UNIVERSITÀ
DI SIENA 1240

UNIVERSITY OF SIENA

DEPARTMENT OF BIOTHECNOLOGY,
CHEMISTRY AND PHARMACY

DOCTORATE IN
CHEMICAL AND PHARMACEUTICAL SCIENCES

XXXIV Cycle

COORDINATOR: Prof. Maurizio Taddei

STRUCTURAL AND FUNCTIONAL STUDIES ON PROTEINS RELATED TO HUMAN DISEASES

SSD: CHIM/03

TUTOR:
Prof. Stefano Mangani

CO-TUTOR:
Prof. Cecilia Pozzi

CANDIDATE
Ludovica Lopresti

ACADEMIC YEAR: 2021/2022

Acknowledgements

At the end of my PhD journey, I would like to thank all the people which share with me this adventure contributing to achieve this milestone. First of all, I would like to express my gratitude to my supervisor Prof. Stefano Mangani for the important opportunity he gave me, for his support, his teaching and his guide. I had the opportunity to learn a lot from his knowledge making my Ph.D. experience always productive and exciting, enriching my cultural background. I would like to give a special thanks to my co-supervisor, Prof. Cecilia Pozzi for her support, patience and the time dedicated me. I am deeply grateful to her for all her precious teachings and advices, which allowed me to grow both personally and professionally. Working with her, I learn a lot and I had the opportunity to carry on my PhD projects in a stimulating and never boring environment. I would like to thank Dr. Manuela Benvenuti, always willing to help me and to encourage in all my three years journey, believing in my abilities every time. I learnt a lot from her crystallographic knowledges and from her life advice. A special thanks to Giusy and Federica to share with me this constructive and important experience, for sharing stimulating ideas with me, and also funny and carefree moments. I would like to thank Prof. Maria Paola Costi and her research group of University of Modena and Reggio Emilia. I would like to thank her for kindness and for her time. From our motivating collaboration, precious for my PhD path, I had the opportunity to enrich my knowledge and improve my work. A particular thank goes also to my family, my boyfriend Nicola and my friends, without their great understanding and encouragement in the past few years, it would be impossible for me to complete my path.

Ludovica Lopresti

Abstract

Protein-protein interactions (PPIs) represent a vast and complex network of intermolecular relationships essential for organism, in which they control a wide range of biological processes. This network plays a vital role in the initiation and progression of cancer; hence, its understanding is crucial for identifying the key functional modulators of tumor progression and metastasis, and for the therapeutic intervention. During my PhD, I had the opportunity to contribute to the structural and functional studies on two different PPIs related to human cancer, the human thymidylate synthase (hTS) homodimer and the heterodimeric interaction between human yes-associated protein (hYAP) and human transcriptional enhanced associate domain 4 (hTEAD4). In the first chapter, the monomer-monomer interface of hTS was exploited to determine and deeply investigate residues critical for the quaternary assembly of the functional enzyme, employing a mutagenic approach. Hence, two sets of hTS interface variants, defined alanine and charged interface variants, have been generated and investigated through circular dichroism (CD) thermal denaturation, kinetic analyses, and X-ray crystallography. We have studied the two interface hot spots F59 and Y202, for which previous studies demonstrated the importance in the hTS dimerization process. The hTS variants F59A and Y202A have been analyzed by means of their thermal stability profiles in comparison to the native enzyme, and by attempting/performing their structural characterization. Notably, the crystallographic structure of the Y202A mutant has been also determined in complex with a dimer disrupter inhibitor, here named compound 1, providing significant information for the rational design of novel hTS interface-targeting molecules. In the second approach, we pointed the attention on the interface residues Gln62 and Thr251 facing themselves on the dimer halves, which have been mutated into Arg and Glu, respectively. The introduction of these charged amino acidic residues affects both the enzyme activity and stability. Our mutagenic approach successfully led to the destabilization of the hTS homodimer, as shown by their structural analyses, evidencing a slight aperture of their quaternary assembly, localized in the interface areas surrounding the mutated residues, also perturbing the active/inactive conformational equilibrium of the enzyme. The interactions of the new variants with the physiological substrate dUMP and its analogue FdUMP have been further characterized. Thus, our hTS interface variants, having a more accessible monomer-monomer interface area may represent useful tools exploitable to identify and screen innovative interface-targeting inhibitors. In the second chapter, the heterodimeric interaction between the Hippo-pathway terminal effector proteins, hYAP and hTEAD4, was examined. To expand the limited knowledges on this key PPI for tumorigenesis and on its modulation, we developed protocols for co-expression, co-purification and crystallization of hYAP TEAD-binding domain (TBD): hTEAD4 YAP-binding domain (YBD) complexes (named S, M, and L complexes). By means of X-ray crystallography, the structure of complex S has been solved, defining new targetable interfaces between the partner proteins. Notably, this work reports the development of the first protocol for the heterologous co-expression and co-purification of full-length hYAP in complex with hTEAD4-YBD (named XL complex), as well as its preliminary structural characterizations. The pilot data here obtained on the hYAP:hTEAD4 complex XL proposes the formation of extended interface areas, larger than those occurring in the complex with the hYAP TBD fragment. The study of these complexes provides interesting new clues helpful for the rational design and the development of molecules able to modulate this PPI playing a critical role in wide-range forms of cancer.

Contents

III Structural insight into hYAP1-hTEAD4 protein-protein interaction: a novel target for cancer treatment 122

List of Abbreviations

5−FTHT	5−formyl−tetrahydrofolate di−L−glutamate
5FU	5−fluorouracile
5−HMTHF	5−hydroxymethyl−6−tetrahydrofolate
aa	amino acid
ASU	Asymmetric unit
Bio SAXS	Biological Small Angle X−ray Scattering
bTS	Bacterial thymidylate synthase
CD	Circular dichroism
CME	S,S−(2−hydroxyethyl)thiocysteine
DHF	7,8−dihydrofolate
DHFR	Dihydrofolate reductase
DLS	Diamond Light Source
DNA	Deoxyribonucleic acid
DNA−BD	DNA Binding Domain
dNTP	Deoxynucleotide triphosphate
dTDP	Deoxythymidine−diphosphate
dTMP	Deoxythymidine monophosphate
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
dUDP	Deoxyuridine diphosphae
dUMP	Deoxyuridine monophosphate
dUTP	Deoxyuridine triphosphae
E. Coli	*Escherichia coli*
***Ef*TS**	*Enterococcus faecalis* TS
EM	Electron microscopy
ESI−MS	electrospray ionization mass spectrometry
ESRF	European Synchrotron Radiation Facility
FAD	Flavin adenine dinucleotide
FDTS	Flavin−dependent thymidylate synthases
FdUMP	Fluorodeoxyuridine−monophosphate

FdUTP	Fluorodeoxyuridine triphosphate
FPGS	Folylpoly−γ−glutamate synthetase
FRET	Fluorescence resonance energy transfer
GARFTase	Glycinamide ribonucleotide formyltransferase
GST	Glutathione−S−Transferase
H4F	Tetrahydrofolate
HEPES	N−2−Hydroxyethylpiperazine−N'−2−Ethanesulfonic Acid
HNSCC	Head and neck squamous cell carcinoma
hTEAD1−4	Human Transcriptional enhanced associate domain
hTS	Human thymidylate synthase
hYAP	Human Yes associated protein
IC50	Half maximal inhibitory concentration
IPTG	Isopropyl β−D−1−thiogalactopyranoside
ITC	Isothermal titration calorimetry
LATS1−2	Large Tumour Suppressor Kinases
LB	Luria Bertani
LC−MS	Liquid Chromatography−Mass Spectrometry
MALDI	Matrix assisted laser desorption/ionization
MBP	Maltose Binding Protein
MDM2	Murine double minute 2
MOB1	Mps1−binder−related 1
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MST1−2	Mammalian Ste20−like 1 and 2
mTHF	N5,10−methylene−5,6,7,8,−tetrahydrofolate
MTX	Methotrexate
MW	Molecular weight
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NCAIS	Non−classical antifolates inhibitor
NCS	Non−Crystallographic Symmetry
NMR	Nuclear magnetic resonance
NR	Non−regular

OD_{600nm}	Optical density at 600 nm
OS	Osteosarcoma
pABA	p−aminobenzoic acid
PAGE	Polyacrylamide gel electrophoresis
PATs	Palmitoyl acyltransferases
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PDX	Pralatrexate
PEG	Polyethylene glycol
PMSF	Phenylmethylsulfonyl fluoride
PMX	Pemetrexed
PPI	Protein−protein interaction
RCC	Renal Cell Carcinoma
RFC	Reduced folate carrier
RNA	Ribonucleic Acid
RTX	Raltitrexed
SAXS	Small Angle X−ray scattering
SB	Super Broth
SDS−PAGE	Sodium dodecyl−sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SHMT	Serine hydroxymethyl transferase
SPR	Surface plasmon resonance
TAD	Transcription activation domain
TAZ	WW domain–containing transcription regulator 1
TBD	TEAD−Binding Domain
TEAD	Transcriptional enhanced associate domain
THF	Tetrahydrofolate
Tt_m	Midpoint of the unfolding transition
TOF	Time of flight
TRIS	tris(hydroxymethyl)aminomethane
TS	Thymidylate synthase
UV	Ultraviolet

VGLL	Vestigial–like
VP	Verteporfin
WB	Western blot
Wt	wild type
YAP	Yes–associated protein
YBD	YAP/TAZ binding domain
ZYP–5052	ZYP–5052 auto–induction medium

Part I

General Overview

1.1 Protein-protein interactions (PPIs): basic concepts, implication in cancer and main experimental applied in PPI studies

Multi-protein complexes are essential biomolecular machines present in all kingdoms of life. They are involved in almost all biological reactions, being crucial to life itself and to the pharmaceutical industry as therapeutic targets and agents. Within a cell, a great number of proteins works cooperatively to make specific functions, mainly performed by establishing intermolecular interactions with other proteins, or with DNA and RNA [1]. Collectively, cellular biomolecular interactions constitute the *"interactome"*, currently a key topic in systems biology [2].

In living organisms, protein-protein interactions (PPIs) are a vast and complex network of intermolecular relationships essential for performing and regulating most biological processes [3], including catalysis of metabolic reactions, DNA replication and transcription, intracellular molecular transport, formation of immunocomplexes and signal transduction. It has been highlighted that over 80% of proteins do not work alone, but in complexes with partner proteins [4]. Proteins involved in the same cellular processes across species are often conserved, together with their interaction network [5]. The same protein can interact with several different partners, being involved in the regulation of a plethora of processes. Thus, PPIs are central hubs for various intracellular signaling pathways, defining a complex cellular scenario. Proteins acting as enzymes, adaptors, transcriptional factors, co-transcriptional regulators, and intrinsically disordered proteins participate to a complex and large set of interactions [6]. Establishing different PPIs can lead to different functional outcomes, e.g. by modulating kinetic properties, substrate specificity and channeling or conformational changes of an enzyme. Through the activation/deactivation of specific partners, PPIs can have regulatory roles at either the upstream or the downstream level of a pathway [7].

A PPI is defined as "direct" if the molecular interfaces of the two partner proteins are in physical contact. Otherwise, if their interaction in the complex is mediated by other proteins, the PPI is defined as “indirect” or “functional”. Here, only direct binary PPIs are discussed. The binary PPIs can be variably classified according to the structural and functional characteristics of the partner proteins. First, they may occur between two or more identical (homo-oligomers) or different (hetero-oligomers) protein units [8]. Homo-oligomers can be further classified as isologues or heterologous [9]. In binary isologues assemblies, the interaction between protomers is mediated by the same surface on both units. On the other hand, heterologous assemblies occur through different interfaces, since the interactors differ as well [10]. Second, PPIs can be defined as obligate or non-obligate. Homo-oligomers are typically characterized by obligate PPIs, and they function only upon complex formation. In contrast, hetero-oligomers are usually composed of units that can exist in cells also in the isolated state [10]. Third, considering the persistence of the interaction, PPIs are divided into transient or permanent. The formers are PPIs that quickly form and dissociate under certain biological conditions, e.g. allowing cells to quickly respond to external stimuli. Transient PPIs commonly involve globular domains and small structural epitopes, like short peptide motifs [11], and they control many cellular processes, like signal transduction, cell growth, metabolic function, DNA replication and transcription. It is well known that the recruitment and the assembly of transcription complexes at specific promoters, as well as the interaction of transcription factors with either DNA cis-elements or co-transcriptional regulators are transient. On the other hand, permanent complexes persist for long time and are mediated by extended and stable interface interactions on partner proteins [11]. Obligate interactions are usually permanent, while non-obligate interactions can be transient or permanent [4, 10]. This classification is not absolute, and the nature of protein-protein associations can be affected by environmental conditions.

The protomers’ concentration and their free binding energy drive the complex formation; notwithstanding the availability of the interacting interfaces, acting as binding sites on the partner proteins, is a fundamental requisite [12]. Specific features of PPI interfaces are hydrophobicity and flatness [13]. Even thought, PPI surfaces can be very large [2] , these surfaces are characterized by the presence of *hot spots*, which drive the recognition between the protomers, contributing to a significant portion of the overall interaction binding energy. A *hot spot* is defined as a critical residue that, if mutated to an alanine, leads to a difference in binding energy >2 kcal/mol [14]. Determination of *hot spots* is crucial to unveil the mechanism of complex formation and their mutation significantly alters complex formation and stability [15]. The experimental identification of *hot spots* is achieved through alanine scanning mutagenesis studies. In this approach, the residue of interest is mutated into an alanine, thus the wild type and mutated proteins are analyzed to determine differences in their free energy of binding [16]. Alternatively, the effects on PPI stability can also be investigated by mutating the *hot spot* into other residues that perturb the interactions between the partner interfaces. The identification of PPI *hot spots* gives important insight into the mechanism of complex formation, exploitable for drug discovery purposes. Alterations of PPI *hot spots* are also at the basis of several types of cancer, where these mutations perturb the formation of physiological protein complexes, leading to pathological outcomes. The mutagenetic approach (to alanine or other amino acid) can thus provide key information on role of *hot spot* residues and the biological effects induced by their modification [17].

The composition analysis of PPI interfaces has revealed that these surface areas are generally enriched in both aromatic and aliphatic hydrophobic amino acids, but also arginine and asparagine are often present [18]. Tryptophan, arginine and tyrosine are the most common *hot spots* found in interface regions [19]. As matter of fact, hydrophobic interactions drive PPI formation, but hydrogen bonds and salt bridges mainly contribute to both binding specificity and stabilization [8]. In some protein complexes, the macromolecular units are covalently linked by disulfide bridges, giving additional structural stabilization to these oligomers [18]. For instance, insulin is composed by two chains linked together by disulfide bonds [20]. Moreover, collapsin-1, a member of the semaphorin family of axon guidance proteins, covalently dimerizes, a requirement to perform collapse activity [21]. Also, fibronectin is a protein dimer comprised of two covalently linked antiparallel units [22]. Different secondary structure elements can be observed at the interface areas; α-helices and β-strands are the main components of homodimer interfaces, nevertheless non-regular (NR) structures, such as turns and loops, principally contribute to intersubunit interactions in heterocomplexes [23]. Based on secondary structure composition, several classes of interfaces can be distinguished: $\alpha-\alpha$, $\beta-\beta$, mixed $\alpha-\beta$, $\alpha-$NR, $\beta-$NR and NR$-$NR [23].

To date, the human interactome consists of over 64,000 binary PPIs (Interactome Atlas) [24, 11]. The elucidation of the whole human cellular interactome remains challenging, but it is fundamental to explain the biological processes occurring under physiological and pathological conditions. Alterations of PPIs are at the basis of a large number of human pathologies, e.g. cancer, infectious and neurodegenerative diseases, thus PPI targeting represents a valid therapeutic strategy [4, 25]. In cancer, PPIs endorse tumorigenesis, tumor progression, invasion, and metastasis; furthermore, they also regulate the biological mechanisms accounting for the development of the physical and molecular features defined as hallmarks of cancer. For instance, PPI alterations are responsible for resistance to cell death, proliferative signaling, promotion of inflammation, induction of angiogenesis, invasion, and metastatic processes. Moreover, some PPIs contribute to the development of multiple oncogenic features; it is well known that MDM2–p53 and Myc–Max play key roles in evading growth suppression and cell death, as well as in promoting genomic instability and cancer metabolism [13].

Aberrant PPIs have been observed in both solid tumors and leukemia, such as breast and prostatic cancer, metastatic melanoma, multiple myeloma, chronic lymphocytic and acute myeloid leukemia [26]. For example, in chronic lymphocytic leukemia, the antiapoptotic proteins Bcl2 and Bcl-xL prevent the physiological oligomerization between Bak and Bax pro-apoptotic proteins and the subsequent apoptosome formation, resulting in the extended survival of leukemic B cells [27]. Other than promoting tumorigenesis, PPI alterations are also involved in the cellular processes leading to the development of drug resistance in cancer cells. The study of PPI relationships in cancer biology has significantly increased in last years, and their targeting represents a promising anticancer approach [13]. Targeting PPIs is a challenging drug discovery process since the peculiar features of these interacting interfaces [26]. Indeed, these areas are large and mostly hydrophobic with quite unique shapes, often flat and few grooves and pockets; these peculiar features complicate the design of targeting molecules. PPI modulators should thus be able to cover large surface areas and to take hydrophobic contacts, often leading to high molecular weight and poorly soluble compounds that would compromise their pharmacokinetic profile [28]. PPI targeting is further complicated by the absence of natural ligands, exploitable as starting point for hit identification, and by the poor accessibility of targetable areas, often hidden by the protomers, especially in obligate oligomers [26]. According to the biological role of the targeted interaction, its modulation can be achieved by either PPI inhibitors or stabilizers. These modulators can bind directly at the protein-protein interface (orthosteric modulation) or indirectly to allosteric sites (allosteric modulation). For example, the small molecules RG7112 and RO5045337 are orthosteric inhibitors of MDM2-p53 interaction currently in clinical phase; whereas Eribulin mesylate and Vinflunine detartrate are allosteric inhibitors of Tubulin-α– tubulin-β complex, used in the treatment of breast and bladder cancers, respectively. In the past years, several anticancer compounds targeting PPIs have been developed and, now, they are at different stages of clinical trials [28]. On the chemical standpoint, PPI modulators are classified into three groups: small molecules, antibodies, and peptides. The first type is often applied to target tight, narrow areas, surrounding specific PPI *hot spots* [26]. To cover huge PPI interfaces, antibodies and peptides are usually more effective. Although monoclonal antibodies selectively recognize and bind to large interface areas and are widely used in clinical treatments, their practice is expensive and often leads to adverse immune reactions. On the other hand, PPI-targeting peptides are designed to target *hot spot* areas of one protomer, mimicking the structure of the partner protein. These peptides have high specificity and affinity for the protein target, but the degradation by cellular hydrolases could limit their efficacy [26, 29]. The investigation of PPIs cannot rely on a single technique but it combines different approaches, including *in vitro*, *in vivo* and *in silico* methods in a complementary fashion [6]. The main *in vitro* approaches used for this purpose are affinity chromatography, co-immunoprecipitation, mass spectrometry (MS), ultracentrifugation, nuclear magnetic resonance (NMR), isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), circular dichroism (CD) spectroscopy , X-ray crystallography and electron microscopy (EM) [1, 5, 11, 28, 30].

Structural biology, collectively including information obtained through X-ray crystallography, single particle Cryo-EM and NMR studies, allows to characterize the molecular basis of PPIs providing essential knowledge on their mechanisms of formation and regulation, exploitable for the development of effective modulators [31]. Additional low-resolution structural information can also be achieved by Small Angle X-ray scattering (SAXS) studies, allowing to determine the size of protein complexes and, hopefully, to rebuild their external shape. Among these techniques, X-ray crystallography represents the 'gold standard' for the atomic-level characterization of macromolecule structures and PPI complexes [32]. Indeed, through crystallographic studies we can provide atomic-level views of the interacting surfaces in macromolecular complexes. To date, several biological complexes, playing pivotal cellular roles, have been characterized through X-ray crystallography; meaningful examples include eukaryotic and prokaryotic ribosomes and their functional complexes [33, 34, 35, 36], the proteasome [37, 38, 39], and RNA polymerase [40, 41, 42]. This technique allows also high-resolution characterization of drug binding to PPI interfaces [28], elucidating their binding modes and giving crucial information to support rational design of more powerful modulators [26].

1.2 Overview of the thesis work

During the three-years of my Ph.D., carried out in the laboratory of Prof. Stefano Mangani and Prof. Cecilia Pozzi, my research activities have been focused on structural and functional studies of protein-protein complexes related to human diseases and, mainly, cancer. The aim of this work was to gain new mechanistic insights into selected targets and their PPIs, supporting the identification and development of innovative inhibitors. Through X-ray crystallography and other biophysical techniques, two different PPIs have been investigated with the goal to obtain valuable data about the mechanistic aspects of the protein inhibition to progress towards more knowledge in the medicinal chemistry field and eventually to design and synthetize better inhibitors against the here studied PPIs. Using X-ray crystallography, I have been able to obtain atomic-level information on protein-ligand and protein-inhibitor complexes. During the development of this Ph.D. thesis project, my activities spanned from molecular cloning to the development of methods for heterologous protein expression in prokaryotic systems and of reliable protocols for their purification, and to the characterization of proteins by means of biochemical, spectroscopic and kinetic analyses and X-ray crystallography.

In the **second part**, I describe the studies carried out on human thymidylate synthase (hTS), to target both the enzyme active site and the monomer-monomer interface of the hTS homodimer (Figure 1.1). TS is an obligate homodimeric enzyme which provides the only *de novo* source of dTMP required for DNA biosynthesis, thus it represents an important target in the oncology field [43]. Several chemotherapeutic agents targeting the hTS active site are currently in clinical use, but they are characterized by limited efficacy for the onset of drug resistance due to TS overexpression. Indeed, hTS acts as translation repressor of its own mRNA (TS-mRNA). Despite the TS-mRNA binding site on hTS is still uncharacterized, the experimental evidence suggests the implication of the hTS dimer interface in hTS-mRNA recognition and binding. Moreover, since hTS is an obligate homodimer, the alteration or disruption of its dimer assembly inactivates the enzyme. Thus, among the new strategies to block hTS activity without triggering drug resistance mechanisms, there is the disruption of its homodimer quaternary assembly. For this purpose, in this work, we combine targeted mutagenesis, structural and biophysical studies to understand the role of important interface residues in the molecular recognition of the partner subunit of hTS, providing key information for the rational design of interface-targeting inhibitors. Here, mutagenic studies at the hTS homodimer interface have been performed to deeply investigate the role of few identified *hot spots* in the enzyme dimerization process. Two sets of hTS interface variants, alanine and charged interface variants, have been generated and investigated by means of circular dichroism (CD) thermal denaturation and kinetic analyses, and X-ray crystallography. The introduction of charged amino acid residues in positions 62 and 251 effectively led to the destabilization of the hTS homodimer, perturbing also the conformational equilibrium between the active and inactive states of the enzyme. The interactions of the new variants with the physiological substrate dUMP and its analogue FdUMP have been further explored. On the other hand, the alanine mutants are more similar to the native enzyme, resembling its structural and functional properties. The determination of the X-ray structure of the hTS Y202A mutant in complex with an interface-targeting inhibitor has provided meaningful information for the rational design of novel hTS interface-targeting molecules. Thus, the findings hither reported show new clues on the mechanism regulating this obligate homo-oligomeric enzyme, and provide insights for its pharmacological inhibition. In the **third part**, I describe the studies performed on two proteins belonging to the Hippo pathway, namely hYAP (human Yes-associated proteins) and hTEAD4 (human Transcriptional Enhanced Associate Domain 4), with a special focus on their heterodimeric interaction. hTEAD4 is a transcriptional factor able to induce gene transcription only upon hYAP binding. Under pathological conditions, dysregulations of hYAP:hTEAD4 functional association can lead to tumorigenesis, making this PPI a target for the development of innovative anticancer drugs [44]. Since its recent discovery, targeting YAP:TEAD interaction has been raised as an attractive therapeutic strategy in the oncology field. To date, structural information about hYAP and hTEAD4 full length proteins is missing and knowledge about their interaction is incomplete. The restricted structural data available on this PPI limit the development of hYAP:hTEAD4 inhibitors. The design of new molecules targeting this PPI remains challenging, but of great interest for the scientific community. To explore this PPI and its modulation, we have focused on the characterization of the YAP-Binding Domain (YBD) of hTEAD4 and of its complex with either hYAP or different segments of its TEAD-Binding Domain (TBD) (Figure 1.1). On this purpose, I have developed protocols for co-expression, co-purification and crystallization of hYAP TBD:hTEAD4 YBD complexes. The structural results achieved on these complexes have expanded the current knowledge on this PPI, defining new targetable interfaces between the two protein partners. Furthermore, this work reports the development of the first protocol for the heterologous co-expression and co-purification of full-length hYAP in complex with hTEAD4-YBD (Figure 1.1). The preliminary data obtained on the hYAP:hTEAD4 complex suggests the formation of interface areas larger than those occurring in the complex with the isolated hYAP TBD. Progress in the study of this complex will provide key information on the interactions between these human oncoproteins. At the end of this manuscript, the list of published articles, posters and conference communications, done during my Ph.D. program, is also reported.

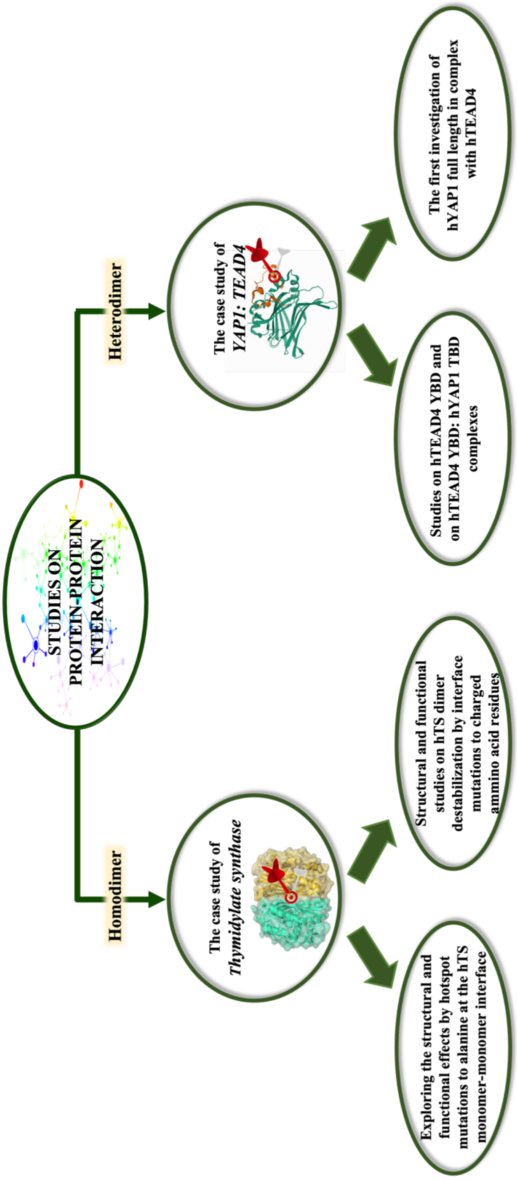


Figure 1.1: Schematic overview of the main studies reported in the thesis.

Part II

Structural and functional analysis of homodimeric enzymes: the case study of Thymidylate Synthase

Chapter 2

Introduction

Thymidylate synthase (TS, E.C. 2.1.1.45) is a highly conserved enzyme, which has a pivotal role for cell survival and replication, since its implication in DNA biosynthesis [45, 46]. TS is part of the folate metabolic pathway, where it accounts for the *de novo* pyrimidine biosynthesis (Figure 2.1 A) [43, 46, 47]. This section is focused on human TS (hTS), representing a validated target for anti-cancer chemotherapy. Various hTS inhibitors, targeting the enzyme active site, are currently applied in clinic; nonetheless, their efficacy is limited by the onset of drug resistance due to TS overexpression [42, 47, 48, 49]. A deeper understanding of the drug resistance mechanisms involving this pivotal enzyme would allow the development of improved anticancer treatments. This section resumes the current knowledge on the catalytic mechanism, the translation regulatory role, and the structure of hTS, and the main inhibition strategies reported in literature.

2.1 Enzymatic role of hTS

hTS catalyzes the reductive methylation of 2'-deoxyuridine-5'-monophosphate (dUMP) to 2'-deoxythymidine-5'-monophosphate (dTMP) using the cofactor N5,10-methylene-5,6,7,8,-tetrahydrofolate (mTHF), as reductant agent and carbon donor (Figure 2.1 B) [46]. At the beginning of the catalytic process, the thiolate moiety of active site Cys195 attacks the carbon atom in position 6 (C6) of the dUMP pyrimidine base, forming a covalent adduct. This bond activates the nearby C5 atom of dUMP, which covalently binds the cofactor mTHF, forming a ternary complex. At this stage, the dUMP C5 accepts the methyl moiety and the hydride donated by the mTHF C11, leading to the formation of the products dTMP and 7,8-dihydrofolate (DHF), afterwards released from the enzyme [46, 52, 53]. Inside cells, the cofactor mTHF is cyclically regenerated by the subsequent activities of two enzymes, dihydrofolate reductase (DHFR), which uses NADPH to reduce DHF to 5,6,7,8-tetrahydrofolate (THF), and serine hydroxymethyl transferase (SHMT), that methylates THF to mTHF, concomitantly converting serine to glycine (Figure 2.1 B) [46].On the other hand, the TS product dTMP is then phosphorylated by thymidylate kinase to 2'-deoxythymidine-diphosphate (dTDP), which is subsequently phosphorylated

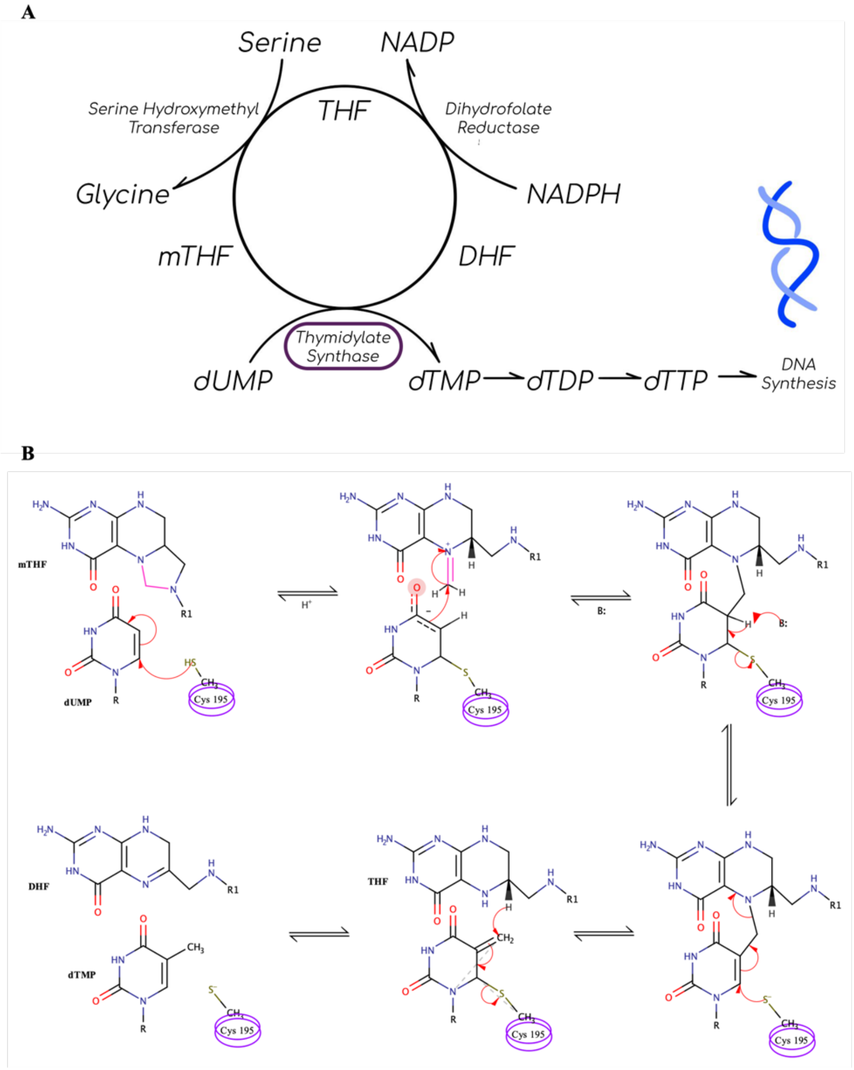


Figure 2.1: (**A**) Schematic diagram of de novo d TMP synthesis involving hTS. (**B**)Scheme of the reaction mechanism proposed for the TS catalyzed reaction [50, 46, 51].

by nucleoside kinases to 2'-deoxythymidine-triphosphate (dTTP), and thus incorporated during DNA biosynthesis [46, 54]. In human cells, hTS provides the only *de novo* synthetic source of dTMP. The inactivation or downregulation of hTS alters replication and leads to apoptosis of rapidly dividing cells, a process known as "thymineless death" [55, 56]. This makes hTS an important target for anticancer therapy (*vide infra, section 2.5*) [43, 47, 49, 57].

2.2 Structure of hTS

hTS is an obligate homodimer, having each subunit composed of 313 residues (molecular weight of $\approx$ 35 kDa). Several hTS structures have been deposited in the Protein Data Bank (PDB) over the last years, showing the enzyme in its apo-state and in different binary and ternary complexes with substrates and inhibitors [47]. hTS has a mixed α/β-fold, showing 7 α-helices and 10 β-strands arranged in three layers: a six-stranded β-sheet, forming the monomer–monomer interface, a long α-helix spanning across the β-sheet and flanked by two shorter helices, and a mixed layer containing the other four helices and two antiparallel two-stranded β-sheets [43, 47]. In addition, each subunit contains two domains: the large (residues 1–98 and 130–313) and the small (residues 99–129) domains. The former domain consists of five α-helices and the six-stranded β-sheet, whereas the latter domain of four α-helices [43, 58, 59]. hTS works as obligate homodimer, indeed residues from both subunits contribute to the active site of each protomer [53, 60, 61]. The hTS active site is a deep L-shaped cavity, characterized by two pockets hosting the substrate dUMP (substrate binding pocket) and the cofactor mTHF (cofactor binding pocket) [60, 62]. The structural information available on hTS, highlights the occurrence of different functional enzyme states, named active and inactive conformations [59, 63]. From a functional standpoint, these conformations of hTS homodimers, mainly differ in the orientation of the catalytic loop, including residues 181–197 and bearing the catalytic Cys195. In the active conformation, Cys195 is exposed inside the catalytic cavity and the enzyme can interact with its physiological substrates dUMP. On the other hand, in the

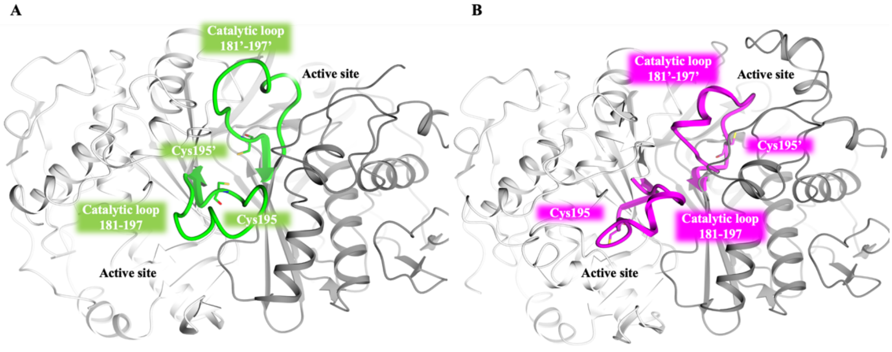


Figure 2.2: Cartoon representations of inactive (**A**) and active (**B**) conformations of hTS (in white cartoon), differing for the orientation of the catalytic loop, residues 181-197 (in green and magenta cartoon, respectively). The position of the catalytic Cys195 (in sticks, green and magenta carbon atoms in the inactive and active conformation, respectively).

inactive conformation the catalytic loop is twisted by ≈180° shifting Cys195 outside the catalytic cavity, toward the dimer interface [53, 60, 64]. Other structural differences are localized in regions 99–129 and 144–158, both highly flexible in the inactive conformation (usually missing in structural models), but fully ordered in the active state [61, 65, 66, 67]. A dynamic equilibrium between these functional conformations of hTS exists in solution and within cells. The presence of molecules able to bind within the active site (e.g. substrate, substrate analogs/inhibitors) can shift this equilibrium towards the active conformation, whereas high concentrations of sulfate/phosphate anions have the opposite effect, stabilizing the inactive conformation [58, 64, 67, 61, 68]. Notably, various hTS residues have been shown to play a role in this dynamic equilibrium and their mutation can stabilize the enzyme in a specific conformation. The hTS mutants R163K and V3L are reported to favor the active and inactive conformation, respectively [69, 66].

2.3 hTS dimer interface and monomer-dimer equilibrium

The hTS dimer interface is generated by the interactions between the large six-stranded β-sheets of two facing protomers, thus classifying it as a $\beta - \beta$

interface [43, 58, 70, 23](Figure 2.3). The interface area involves forty-five residues of each monomer [43], and, among them, seven arginines and six aromatic residues [70, 71] are considered pivotal to keep interaction with the cognate subunit [15, 72]. Former computational and experimental studies based on alanine scanning have identified various *hot spots* at the hTS interface (namely Lys47, Phe59, Arg175, Ile178, Trp182, Leu198, Tyr202, and Tyr213) [70, 71]. The interface area is slightly modified by the active/inactive state of hTS but none of these *hot spots* is affected by the enzyme conformation. The modulation of the interface interacting area either by point mutations or by molecules able to bind to it, can interfere with the hTS dimeric assembly, also perturbing the catalytic activity of the enzyme [58, 67, 70, 71]. Further the active/inactive equilibrium, hTS can also shift between the monomeric and dimeric assembly [43, 46, 58]. The monomer is thought to regulate the hTS mRNA translation through a negative feedback mechanism (*vide infra, Section 2.4*). The characterization of the mechanisms regulating this monomer-dimer equilibrium is pivotal to elucidate the hTS regulatory function accounting for drug resistance in cancer cells. Recent evidence on hTS have shown that the monomer-dimer equilibrium is concentration dependent [58, 73]. Ultracentrifugation analysis have shown that at concentration >5 μg mL $^{-1}$ (70 nM) the protein is mainly dimeric, whereas comparable abundances of dissociated and associated hTS dimers are reached at protein concentrations ≈2.5 μg mL^{-1} (35 nM). The abundance of the monomer increases by decreasing the hTS concentration and at concentrations <0.75 μg mL^{-1} the protein is mainly monomeric [73]. Further evidence on this equilibrium have been provided by fluorescence resonance energy transfer (FRET) analysis [63]. By this technique, the equilibrium constant for the dimer dissociation has been determined in solution, resulting of 2 x 10^{-7} M, corresponding to a $\Delta G°$of ≈38 kJ mol^{-1} [63].

2.4 hTS as translational regulator: the hTS-mRNA interaction

hTS is also a mRNA-binding protein, may contributing to the regulation of the expression of various genes; thus, hTS is a critical regulator of several key cellular aspects, including apoptosis and chemosensitivity. *In vitro* studies have indicated that hTS negatively regulates the translation of the textitp53 tumor suppressor gene and of textitc-myc [74], by acting as translational repressor [75, 64, 76, 77]. hTS is also able to bind its own mRNA (hTSmRNA), determining a feedback inhibition mechanism that regulates the intracellular level of hTS [35, 36, 78]. The hTS-hTSmRNA interaction has a high-binding affinity, with a dissociation constant (K_d) in the low nanomolar range [78, 79]. Two different hTS binding sites have been identified on hTSmRNA, recognizing the enzyme by one or both *cis*-acting elements. The first region (site 1) consists of nucleotides (nt) 81–110, adopting a stem loop structure and including the initiation codon [80]. hTS binding to this element can thus stabilize the hairpin loop making the start codon unavailable for ribosomal recognition. The second *cis*-element (site 2) lies between nt 450-520, in the protein-coding region [74, 76, 81]. *In vitro* and *in vivo* studies have shown that each site could interact independently with hTS, nonetheless the concomitant binding to both elements is required to achieve the full translational autoregulatory effect [47]. The active/inactive and monomer/ dimer equilibria of hTS affects the hTSmRNA binding. Indeed, upon binding to the substrate or to substrate-analogue inhibitors, switching the enzyme to the active conformation, reduced levels of hTS-hTSmRNA are observed [64, 81]. The net effect of ligand binding is thus the abrogation of translational repression, resulting in hTS biosynthesis and increased cellular levels [81], responsible for the onset of drug resistance upon treatment with hTS inhibitors binding to the active enzyme [82, 83].The form in which hTS interacts with hTSmRNA is still controversial; either the inactive dimer or the hTS monomer can be responsible for this interaction. Nonetheless, increasing evidences seem to suggest that is the hTS monomer that binds to hTSmRNA. The hTSmRNA binding site on hTS is yet uncharacterized, but experimental evidences suggest the implication of the hTS dimer interface in hTSmRNA recognition and binding [76, 83, 84]. Peptides matching the TS interface sequences 31–47, 56–72, 131–147, 176–192, and 201–217, were reported to bind to hTSmRNA [76].

2.5 Inhibitors of hTS in anticancer therapy

hTS is a well-established target for anticancer therapy and few hTS inhibitors are currently used in clinical practice for the treatment of the most-difficult-to-treat cancers, such as colorectal, pancreatic and non-small-cell lung cancer [85, 86]. The inhibitors to date developed bind to the hTS the active site, resembling either the substrate dUMP (dUMP-like inhibitors) or the cofactor mTHF (folate-like inhibitors).

The nucleotide analogues 5-fluorouracil (5-FU), Capecitabine, and Tegafur are the most effective hTS-targeting drugs [47]. 5-FU is a prodrug, converted inside cell into three main, active metabolites: fluorodeoxyuridine-monophosphate (FdUMP), acting as hTS inhibitor, fluorodeoxyuridine triphosphate (FdUTP), and fluorouridine triphosphate (FUTP), both targeting DNA/RNA biosynthesis [87]. FdUMP forms a slowly-reversible ternary complex with hTS, leading to its inactivation. Capecitabine and Tegafur are also 5-FU prodrugs, activated inside cells by thymidine phosphorylase and uridine phosphorylase, and by cytochrome P450, respectively [88, 89]. On the other hand, the most effective mTHF-like (or antifolate) drugs are methotrexate (MTX), pemetrexed (PMX), raltitrexed (RTX, also known as Tomudex), and pralatrexate (PDX). The cellular internalization of folate analogues require a carrier-mediated uptake, usually performed by the folate transporter, the reduced folate carrier (RFC) [90, 91]. Moreover, inside cells, they undergo to poly-glutamylation by folylpoly-γ-glutamate synthetase (FPGS) [92, 93]. Cofacor-like drugs are able to act as multi-targeting agent, inhibiting various folate-dependent enzymes, such as dihydrofolate reductase (DHFR) and glycinamide ribonucleotide formyltransferase (GARFTase) [93]. Raltitrexed is a quinazoline folate analogue that selectively inhibits hTS (IC_{50}= 0.38 μM), currently applied for the treatment of various solid tumors, including gastric cancer [94] head and neck cancer [95] malignant mesothelioma [96] as well as colorectal cancer [97]. This drug has been studied in synergistic drugs combinations, aiming to potentiate the activity of both drugs and to limit their toxicity [62]. The raltitrexed/cisplatin combination was shown to be more effective then cisplatin alone, improving the outcome and the life-quality of patients with malignant pleural mesothelioma [96]. A very recently study has provided evidence that sequential administration of RTX followed by 5FU, inhibit ovarian cancer cell growth *in vitro* in a synergic mode. Also, the synergic inhibition of hTS by FdUMP and RTX has also been shown in the structural characterization of their ternary complex (PDB id 6ZXO) [62]. Further the classical mTHF-like drugs, the developed some non-classical antifolates inhibitors (NCAIs) of hTS have been reported [98]. NCAIs are able to passively diffuse into cells, not requiring the active transport by RFC and/or FPGS activity [47, 98]. Plevitrexed and Nolatrexed are two examples of NCAIs [98]102, both acting as effective and selective hTS inhibitors [99, 100].

The main drawbacks of these hTS-targeted therapies are drug toxicity and resistance, multifactorial phenomena due to TS overexpression and impaired function of RFC and FPGS proteins [101, 102, 103, 104]. Drugs targeting the active site stabilize hTS in the active conformation and reduce its affinity for hTSmRNA, thus releasing it from the complex with the enzyme and inducing hTS production [90, 101, 105, 106].

Different inhibition strategies are thus needed to effectively block the enzyme activity without triggering drug resistance phenomena in cancer cells. In 2011, textitCardinale et coworkers identified few octapeptides that inhibit hTS, by binding at the inter-subunit interface. Among them, the LR peptide (LSCQLYQR) was the most promising hTS inhibitor, shown act through an allosteric mechanism. The characterization of its complex with hTS (PDB id 3N5G) demonstrated that it binds to an allosteric site, at the periphery of the dimer interface, stabilizing the inactive conformation of the enzyme. Furthermore, this peptide is effective on resistant ovarian cancer cell lines without inducing hTS overexpression [58]. A very recent study has reported the development of a small molecule, named compound F13

(1-(5-methoxy-1H-indol-3-yl)-N,N-dimethyl-methanamine), able to inhibit hTS by binding to the dimer interface. At variance with the LR peptide, this compound has been shown to interact with the active conformation of the enzyme [67]. The efficacy of the compound on cancer cells has not been reported yet.

2.6 Main approaches applied in this thesis

In this work, we have performed structural and functional analysis of hTS, focusing our attention on the investigation of the interface area. In a first approach, we have studied the two interface *hot spots* F59 and Y202 [70]. The latter residue is located on the fourth strand of the interface β-sheet, and its side chain faces F59' of the partner subunit [47, 70, 71]. Previous studies showed that these residues are important for hTS dimerization, indeed their mutation to alanine induces meaningful destabilization of the dimer assembly. The monomer-dimer equilibrium of these variants was investigated by FRET analysis, reporting a significant gain in their K_d with respect to the wild-type enzyme. In this work, we have deeply investigated the F59A and Y202A mutants, by analyzing their thermal stability profiles in comparison to the wild-type enzyme and by attempting/performing their structural characterization. The structure of the Y202A mutant has been also determined in complex with a dimer disrupter inhibitor, compound 1, developed by the research group of Prof. Maria Paola Costi (University of Modena and Reggio Emilia). Achieving structural information on enzyme-inhibitor complexes is pivotal for the rational design of novel, improved interface binders. In the second approach, we have generated and studied a second set of hTS mutants having interface *hot spot* residues replaced by bulky, charged amino acids. To maximize the destabilization effects induced by these mutations on the enzyme dimer assembly we have chosen residues facing themselves on the dimer halves. To this end, we have focused our attentions on interface residues Gln62 and Thr251, that have been replaced by Arg and Glu, respectively. All interface variants have been characterized by CD thermal denaturation and kinetic

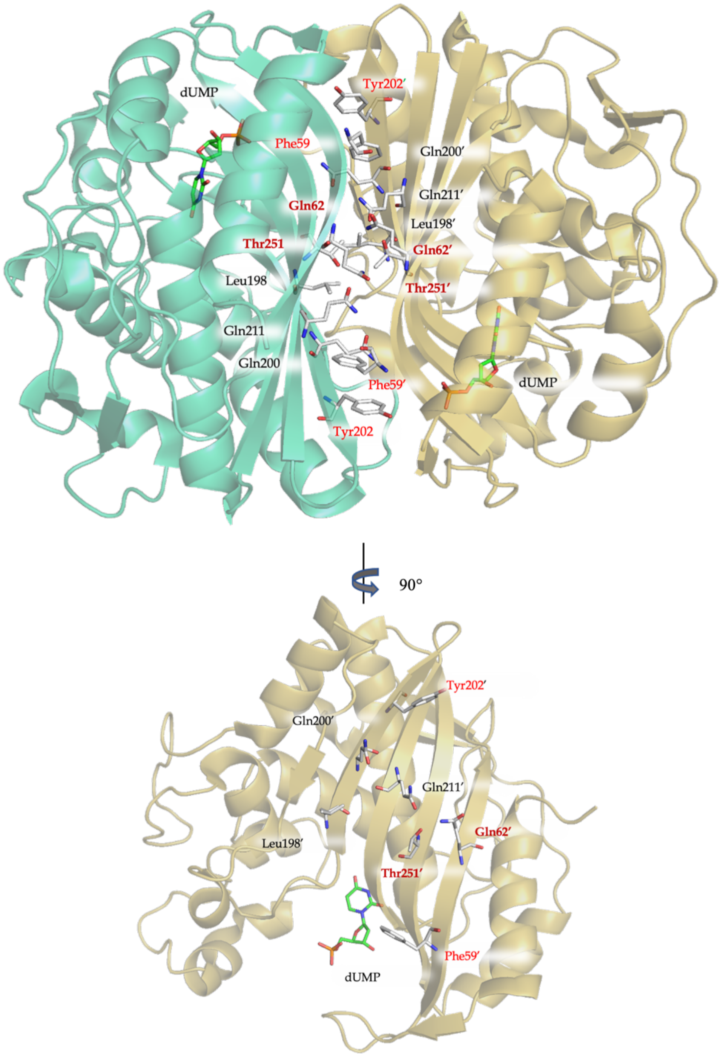


Figure 2.3: Interface view of the hTS homodimer. The mainly residues belonging to the monomer-monomer interface region are displayed in sticks (grey carbon atoms). The interface residues studied in this work (Tyr202, Phe59, Gln62 and Thr251) are highlighted in red. The dUMP molecule (in stick, green carbon atoms) bound the active site is displayed.

analyses, and by X-ray crystallography. The results achieved by their characterization, have provided evidence of the peculiar behavior of these interface variants, showing the meaningful effects induced by the mutations on the enzyme activity and stability. The effective dimer destabilization has also been shown by the structural analysis of the interface variants, evidencing a slight aperture of their quaternary assembly, localized in the interface areas surrounding the mutated residues. Thus, these hTS interface variants, may represent functional tools exploitable to identify innovative interface-targeting inhibitors.

Chapter 3

Exploring the structural and functional effects of hotspot mutations at the hTS monomer-monomer interface

3.1 Materials and Methods

3.1.1 Cloning and Site-direct Mutagenesis

The hTS gene was previously cloned in the pQE-80L vector (Novagen, Darmstadt, Germany), downstream the sequence encoding for a non-cleavable N-terminal His^6-tag, within *BamHI* and *HindIII* restriction sites (plasmid pQE80L-hTS) (Figure 3.1) [72]. The pQE80L-hTS plasmid was kindly provided by the research group of Prof. Maria Paola Costi (University of Modena and Reggio Emilia) and it was used as template to generate alanine mutants. The pQE80L-hTS Y202A plasmid was formerly obtained in the lab according to reported methods [70]. The hTS F59A variant was generated by site-directed mutagenesis using partially overlapping primers (*Table 3.1*), designed according to the protocol defined by *Zheng et al.* [107]. To generate the hTS F59A mutant, the PCR was performed using the AccuPrime Taq DNA Polymerase High Fidelity (Thermo Fisher Scientific, Inc., Waltham, MA, USA), with minor modifications to the manufacturer's protocol. The reaction mixture (50 μL) included 2 ng template DNA (pQE80L-hTS plasmid), 10 μM primer pair, 1 unit (U) AccuPrime Taq DNA Polymerase, and 1X AccuPrime PCR Buffer I. The PCR protocol started with an initial denaturation (94 °C, 30 sec), followed by 25 cycles of denaturation (94 °C, 30 sec), annealing (61 °C, 30 sec), and extension (68 °C, 6 min). Thereafter, an aliquot (10 μL) of the PCR-amplified sample was incubated overnight at 37 °C with 1 μL of Fast Digest DpnI (Thermo Fisher Scientific, Inc., Waltham, MA, USA) to digest the methylated parental DNA template and select for the newly synthesized DNA containing mutations. The resulting sample was heat-shock transformed in chemically competent *E. coli* TOP 10 cells and positive transformants were selected on LB-agar plates supplemented with 100 mg L^{-1} ampicillin. Single colonies were cultured in LB medium (with 100 mg L^{-1} ampicillin) at 37 °C for 16 h, under vigorous aeration (220 rpm). The plasmid was extracted through the E.Z.N.A. ® DNA extraction Kit I (Omega Bio-Tek) and the mutation was confirmed by DNA sequencing (ATAC sequencing, Eurofins Genomics Italy).

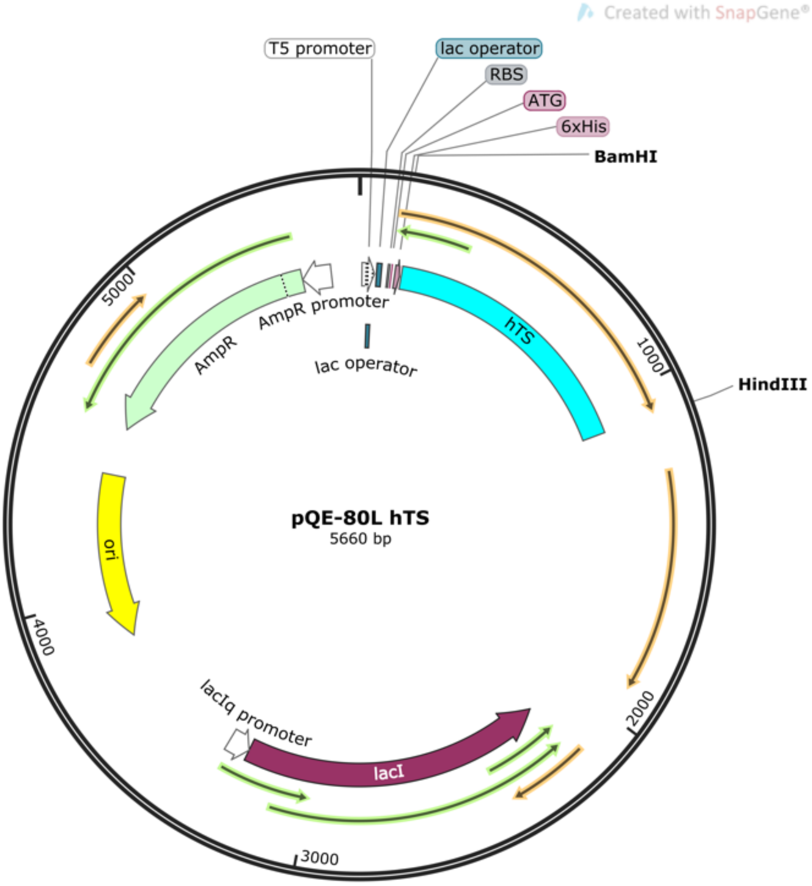


Figure 3.1: Map of the pQE80L-hTS plasmid vector, including the hTS gene cloned within the BamHI/HindIII restriction sites, downstream the sequence encoding for the His^6-tag (indicated as 6xHis in the map). AmpR: Ampicillin resistance gene; lacI: lactose operon repressor; ori: origin of replication.

Gene Target	**Primer**	**Sequence (5'→3')**	**Product**
hTS WT	Fw: F59A	GCGTCGCCGGCATGCAGGCGAG	hTS F59A
	Rv: F59A	CATGCCGGCGACGCTTAAGGTGC	

Table 3.1: Sequences of forward (Fw) and reverse (Rv) primers used for mutagenic PCR to introduce the point mutation F59A. The template used for PCR is the pQE80L-hTS plasmid. The mutated codon on both strands is in red.

3.1.2 Protein expression and purification

Wt His^6-tag hTS (HT-hTS) (recombinant protein sequence in *Table* **??**) was obtained as formerly described [70]. The production of His^6-tag hTS F59A (HT-hTS F59A) and His^6-tag hTS Y202A (HT-hTS Y202A) was carried out following reported protocols [70], with minor modifications. Briefly, the expression plasmids were used to heat-shock transform *E. coli* BL21(DE3) cells and positive transformants were selected on LB-agar plates added by 100 mg L^{-1} ampicillin. Bacterial cells were cultured in LB supplemented with 100 mg L^{-1} ampicillin, at 30 °C for 20 h and protein overexpression was induced at OD_{600nm} of 0.6, adding 0.4 mM IPTG (isopropyl-beta-D-thiogalactopyranoside). Cells, harvested by centrifugation, were resuspended in buffer A (30 mM NaCl and 50 mM HEPES, pH 7.5) added by 20 mM imidazole, 0.2 mM phenylmethylsulfonyl fluoride (PMSF), and 0.5 mg mL^{-1} lysozyme, and disrupted by sonication after 1 h incubation on ice. The soluble cellular fraction was separated from cell debris by centrifugation (13000 ×g, 1 h, 4 °C) and the target protein variants were purified by nickel affinity chromatography (HisTrap FF 5mL column, GE Healthcare), using a three-step concentration gradient of imidazole (80 mM, 250 mM and 500 mM, in buffer A). Fractions containing the target proteins, eluted at 250 mM imidazole concentration, were pooled and extensively dialyzed in buffer A. The high purity (>98%) of the final protein samples was confirmed by Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis and mass spectrometry (MS, *vide infra*).

3.1.3 Protein Characterization: Mass Spectrometry and Circular Dichroism (CD) thermal denaturation analyses

The purified HT-hTS F59A and HT-hTS Y202A were analyzed by ESI-MS (LTQ-Orbitrap mass spectrometer, equipped with an electrospray ionization source interfaced to UPLC Easy LC 1000) on service (Toscana Life Science, Mass Spectrometry facility). Spectra were recorded on water solutions of the protein samples having a concentration of 3 μM. The CD spectra of HT-hTS, HT-hTS F59A and HT-hTS Y202A were recorded at room temperature in the wavelength range 200-250 nm using the Jasco J-815 spectropolarimeter and 0.1-cm quartz cuvettes. Each sample (200 μL) included a 20 μM protein solution in 10 mM HEPES pH 7.5 and 100 mM NaCl (buffer signals were subtracted from sample profiles, during spectra recording). Thermal stability profiles of HT-hTS and the alanine variants were recorded by monitoring the far-UV CD signal at 220 nm, according to published protocols [108]. Protein samples were heated from 25 to 80 °C, at a rate of 2 °C min^{-1}. A second series of spectra was subsequently recorded on all variants incubated with 1 mM dUMP for 30 min on ice. Measurements were performed in triplicates. The midpoint of the unfolding transition (melting temperature, T_m) was thus obtained by analyzing data with the software GraphPad Prism 7 in non-linear regression, using the Boltzmann sigmoidal function.

3.1.4 Crystallization

Crystallization screens were performed for HT-hTS F59A using the commercially available kits JBScreen Basic (JBSB) 1–4 and Classic (JBSC) 6 from Jena Bioscience (Jena, Germany), and PEG/Ion, Index, Crystal Screen 1 and 2, and Grid screen Ammonium Sulfate from Hampton Research (Aliso Viejo, California, USA). Drops, consisting of 2 μL protein solution (20 mg mL^{-1} in buffer A, with and without 20 mM β-mercaptoethanol) and 2 μL precipitant, were equilibrated over 200 μL reservoir. Crystallization screens were performed using the sitting-drop vapor diffusion technique [109], at both 8 °C and 20 °C.

No crystal growth was observed over 1 year of incubation. Crystals of HT-hTS Y202A were obtained by the hanging drop vapor diffusion technique in the so-called high salt conditions, according to reported protocols [70]. Briefly, crystallization drops consisting of equal volumes of protein (20 mg mL_{-1} in 30 mM NaCl, 50 mM N-2-Hydroxyethylpiperazine-N'-2-Ethanesulfonic Acid (HEPES), pH 7.5, and 20 mM β-mercaptoethanol) and precipitant (2.0 M ammonium sulphate, 5% w/v Peg 400, 0.1 M MES, pH 6.5, and 20 μM β-mercaptoethanol) solutions were equilibrated over a 600-800 μL reservoir. Crystals, shown in *Figure 3.2 A*, were observed after few days of incubation at 20 °C. Samples of HT-hTS and HT-hTS Y202A in complex with compound 1 were prepared both by incubation of the concentrated proteins with the inhibitor (concentrated incubation method) and by co-centrifugation of the protein-inhibitor complexes (co-centrifugation method). In the former method, the protein (10 mg mL^{-1} in 30 mM NaCl, 50 mM HEPES, pH 7.5, and 20 mM β-mercaptoethanol) was incubated with compound 1 (5 mM, solubilized in dimethylsulfoxide). On the other hand, in the latter method, diluted protein solutions (1 mg mL^{-1}, in 30 mM NaCl, 50 mM HEPES, pH7.5 and 20 mM β-mercaptoethanol) were incubated with compound 1 (0.5 mM) for 12 h at 4 °C and then concentrated 10-times using centrifugal concentrators having a molecular weight cut-off of 10 kDa (Pierce^{TM} Protein Concentrator PES). The resulting samples were used for crystallization (crystals of Y202A-compound 1 obtained with co-centrifugation method are shown in *Figure 3.2 B*) following the same procedures described for the ligand-free enzymes. The precipitant solution used for HT-hTS crystalizzation was composed of 1 M ammonium sulphate and 0.1 M tris(hydroxymethyl)aminomethane (TRIS) pH 8.5. Prior to data collection, all crystals were washed in cryoprotectant solutions (20% of either ethylene glycol or glycerol added to each precipitant solution) and flash frozen in liquid nitrogen.

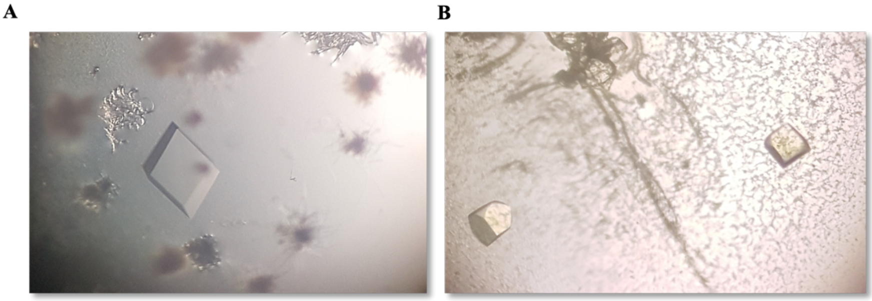


Figure 3.2: Crystals of (**A**) Y202A and (**B**) Y202A in complex with compound 1 (co-centrifugation method).

3.1.5 Data collection, structure solution and refinement

Diffraction data were collected at 100 K, using synchrotron radiation at the Diamond Light Source (DLS, Didcot, UK) beamline I04 equipped with Eiger 2XE 16M detector. Reflection data were indexed and integrated using the program XDS [110] and scaled with SCALA [111, 112] from the CCP4 suite [113]. Data collection and reduction statistics are reported in *Table* **??**. Molecular replacement was performed using the software Molrep [114] and the structure of hTS in the inactive conformation (PDB id 3N5G [58]) as searching model, excluding non-protein atoms and water molecules. The structures were refined through the program REFMAC5 [115] from CCP4 suite [113]. The molecular graphic software Coot [116] was used for manual rebuilding and modeling of the missing atoms in the electron density and to add solvent molecules. Sulfate anions and glycerol and ethylene glycol molecules from crystallization/cryoprotectant solutions were found in our model and modelled according to electron density. Furthermore, the inspection of the Fourier difference map in the structure of the complex HT-hTS Y202A-compound 1 (crystals obtained by the co-centrifugation method) evidenced the presence of the compound, modeled according to the electron density. On the other hand, no additional electron density compatible with the presence of the inhibitor was observed in the structures of the samples obtained by the concentrated incubation method and of the HT-hTS sample obtained by co-centrifugation method. The occupancy of exogenous ligands was adjusted to values resulting

	HT-hTS Y202A	**HT-hTS Y202A** **Compound 1**
Diffraction source	I04 (DLS)	I04 (DLS)
Wavelength (Å)	0.9795	0.9795
Temperature (K)	100	100
Detector	Pilatus 6M-F	Pilatus 6M-F
Crystal-detector distance (mm)	332.90	368.389
Exposure time per image (s)	0.25	0.25
No. of heterodimers in the ASU	1	1
a = b, c (Å)	95.99 95.99 81.97	96.10 96.10 83.40
Resolution range (Å)	58.37-2.4 (2.53-2.40)	58.91-2.55 (2.69-2.55)
Total no. of reflections	176593(25825)	151189 (22393)
Completeness (%)	100.0 (100.0)	100.0 (100.0)
Redundancy	10.1 (10.4)	10.1 (10.5)
$\langle I/\sigma(I) \rangle$	10.4 (2.0)	17.6 (2.4)
CC1/2	0.996 (0.791)	0.999 (0.833)
R_{meas}	0.108 (1.224)	0.089 (1.282)
Overall B factor from Wilson plot (Å^2)	72.7	72.84

Table 3.2: Data collection and processing statistics. Values in parentheses are for the highest resolution shell.

in atomic displacement parameters close to those of neighboring protein atoms in fully occupied sites. The final models were inspected manually, checked with Coot [116] and PROCHECK [117], and finally validated through the PDB deposition tools. Final refinement statistics are reported in *Table 3.2.* Figures were generated using molecular graphic software PyMOL [118] and CCP4 mg [119].

	HT-hTS Y202A	**HT-hTS Y202A Compound 1**
Resolution range (Å)	58.37-2.4 (2.53-2.40)	58.91-2.55 (2.69-2.55)
No. of reflections, working set	16490 (1182)	14137 (1018)
No. of reflections, test set	902 (64)	740 (56)
R_{cryst}	0.198 (0.430)	0.186 (0.407)
R_{free}	0.248 (0.434)	0.248 (0.393)
No. of non-H atoms		
Protein	1813	2127
Ion (Sulfate ion)	10	10
Ligand	-	24
Water	37	65
Other ligants	-	-
Total	1860	2226
R.m.s. deviation Bonds (Å)	0.007	0.007
Angles (°)	1.644	1.700
Average B factors (Å^2)	69.1	74.8
Estimate error on coordinates based on R value (Å)	0.22	0.32
Ramachandran plot Most favored (%)	93.0	90
Allowed (%)	7	10

Table 3.3: Refinement statistics for HT-hTS Y202A and in complex with compound 1. Values in parentheses are for the highest resolution shell.

3.2 Results and Discussions

3.2.1 Expression, purification, and characterization of HT-hTS wild type and its alanine variants, F59A and Y202A

Wt hTS and its interface mutants F59A and Y202A were expressed as His^6-tag recombinant proteins in a heterologous bacterial system, using the BL21 (DE3) *E. coli* strain. HT-hTS and its variants consist of 325 residues, including the 12 amino acids belonging to the unremovable N-terminal His^6-tag (residues numbering refers to the hTS sequence, UniProtKB id P04818) [58]. Analogously to the wt-enzyme, all alanine variants were expressed as soluble proteins and their purification protocol relied on a single-stage nickel affinity chromatography (*Figures 3.3 A and 3.4 A*). The high purity of the protein samples obtained by this procedure was estimated $\geq$ 98% by SDS-PAGE analysis (*Figures 3.3 B and 3.4 B*) and ESI-Orbitrap MS (*Figures 3.5 and 3.6*; HT-hTS F59A: theoretical MW = 37038.4 Da, determined MW = 37049.0 Da; HT-hTS Y202A theoretical MW = 37022.4 Da, determined MW = 37033.0 Da). At variance with the HT-hTS F59A mutants that was obtained in a yield of 20 mg mL^{-1} bacterial culture, the HT-hTS Y202A variant was purified in high yield, estimated to 95 mg L^{-1} bacterial culture, analogously to HT-hTS (120 mg L^{-1}). The residues in positions 59 and 202 are placed at the dimer interface of the enzyme, where they point towards each other. The comparison of the CD spectra recorded for hT-hTS and its alanine variants shows that the overall fold of the protein is not affected by the introduced mutations (data not shown). The CD technique was also used to monitor the thermal unfolding profiles of the proteins and to calculate the midpoint of the unfolding transition (T_m), an effective method to assess the effects of mutations on protein stability [120, 121]. It is worth noting that the curve of HT-hTS shows a biphasic unfolding transition, not detected in the profiles of both hTS F59A and Y202A variants (*Figure 3.7*). Furthermore, the comparison of the thermal denaturation profiles recorded for HT-hTS and the alanine variants shows a slight destabilization of both interface mutants. Indeed, the T_m for HT-

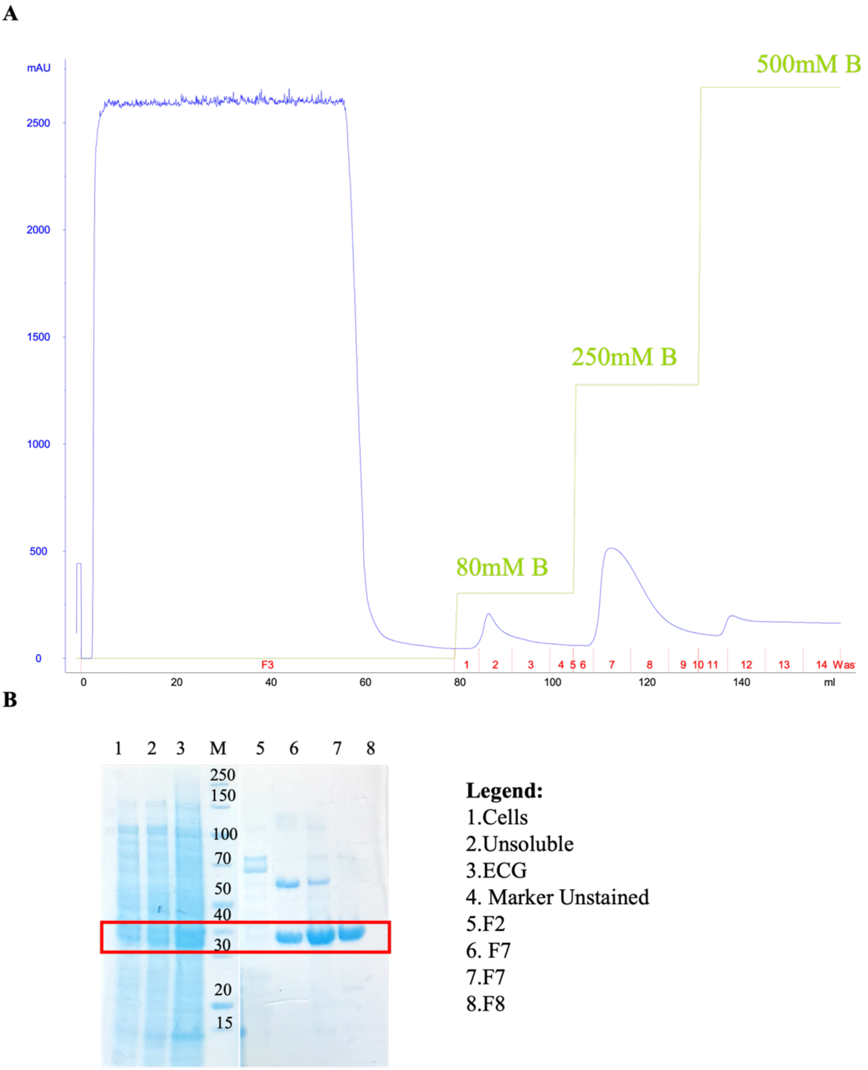


Figure 3.3: (**A**) Chromatogram of the nickel- affinity chromatography (HisTrap FF 5mL column, GE-Healthcare) performed to purify HT-hTS F59A. The UV$_{280nm}$ monitoring and imidazole concentration are shown as blue and green lines, respectively; the elution fractions are indicated on the x-axis as red marks. (**B**) SDS-PAGE analysis of the fractions collected from affinity chromatography (the protein ladder used for MW estimation is in lane 4).

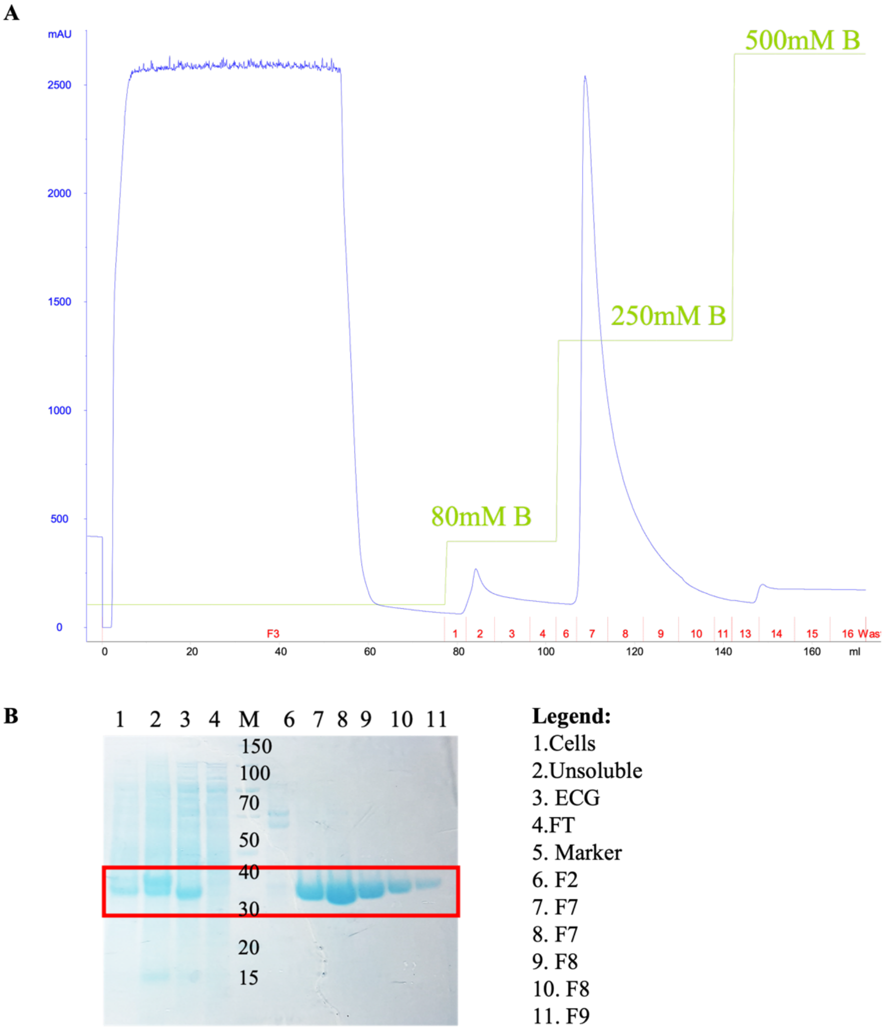


Figure 3.4: (**A**) Chromatogram of the first stage of nickel-affinity chromatography (HisTrap FF 5mL column, GE-Healthcare) performed to purify HT-hTS Y202A. The UV$_{280\text{nm}}$ monitoring and imidazole concentration are shown as blue and green lines, respectively; the elution fractions are indicated on the x-axis as red marks. (**B**) SDS-PAGE analysis of the collected fractions (the protein ladder used for MW estimation is in lane 5).

hTS is 53.26 (± 0.08) °C, compared to T_m of 51.48 (± 0.18) °C and 50.41 (± 0.09) °C for HT-hTS F59A and HT-hTS Y202A, respectively (*Figure 3.7 A*). T_m decreases by ≈ 2 °C and by ≈ 3 °C are observed for HT-hTS F59A and HT-hTS Y202A, respectively, with respect to HT-hTS. Our results are in agreement with former data published on these hTS alanine mutants [70], reporting their reduced dimer stability determined by both computational prediction and FRET analysis. In this work, we have further characterized the effect of the substrate on the protein stability, measuring the unfolding transition curves of HT-hTS and its alanine mutants in presence of dUMP. A stabilization effect is observed for all proteins upon substrate addition, resulting in T_m values of 56.77 (± 0.08) °C for HT-hTS:dUMP, of 51.97 (± 0.16) °C for HT-hTS F59A:dUMP, and of 54.01 (± 0.14) °C for HT-hTS Y202A:dUMP (*Figure 3.7 B*). It is well known that, in presence of dUMP, hTS switches to the active conformation. Our data evidence the improved stability profile of the complex with respect to the ligand-free state of the enzyme, showing a T_m increment of ≈ 3.5 °C. An analogous effect is observed for the HT-hTS Y202A mutant, reporting a T_m increase of ≈

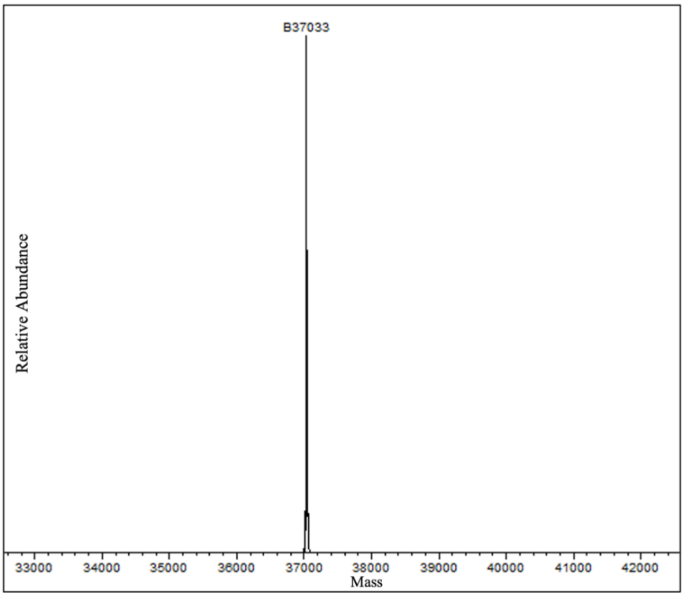


Figure 3.5: ESI-Orbitrap spectrum of HT-hTS F59A (theoretical MW = 37038.4 Da, determined MW = 37049.0 Da).

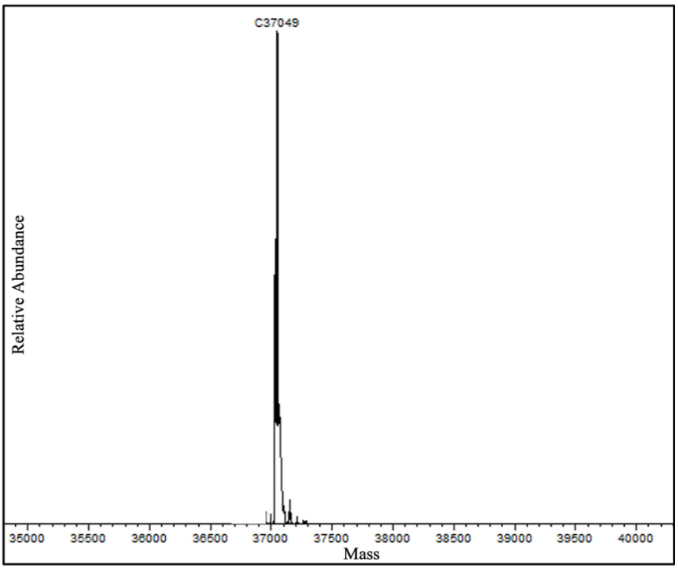
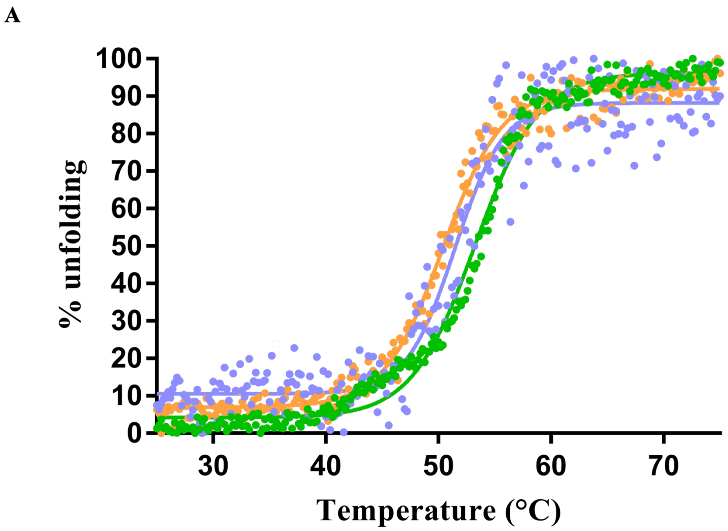
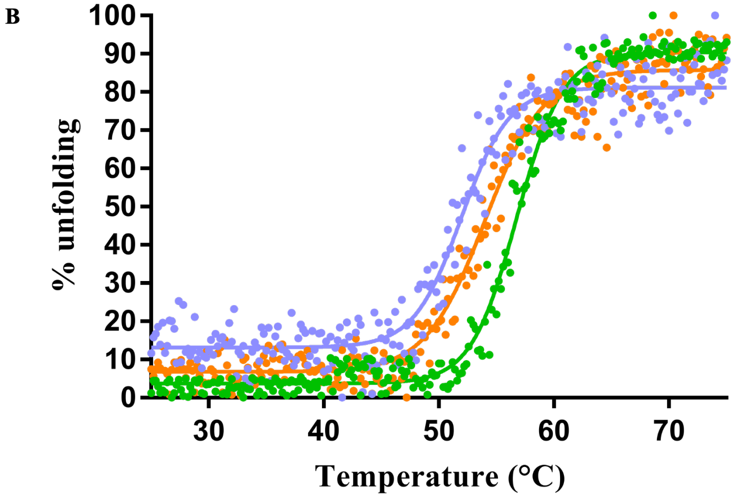


Figure 3.6: ESI-Orbitrap spectrum of HT-hTS Y202A (theoretical MW = 37022.4 Da, determined MW = 37033.0 Da).

4 °C in presence of dUMP. On the other hand, a lower effect is observed for the HT-hTS F59A mutant, displaying a T_m increase of only ≈ 1 °C. This finding correlates with the kinetic profile reported for the mutants [70], showing similar kinetic profiles for HT-hTS and the Y202A variant but altered for the F59A mutant (significantly reduced k_{cat}). Since both mutated residues are not part of the active site, alterations in the catalytic activity and in protein stability should arise from their effects at the hTS protein-protein interface. The results of former and present characterizations of these alanine variants suggest that the mutation F59A induces stronger effects on the enzyme, reporting alterations on both activity [70] and conformational stability.



Protein sample	HT-hTS	HT-hTS F59A	HT-hTS Y202A
Melting Temperature (Tm)°C	53.26 ± 0.08 °C	51.48 ± 0.18 °C	50.41 ± 0.09 °C



Protein sample	HT-hTS	HT-hTS F59A	HT-hTS Y202A
Melting Temperature (Tm) °C	56.77 ± 0.08 °C	51.97 ± 0.16 °C	54.01 ± 0.14 °C

Figure 3.7: Thermal unfolding transition curves of (**A**) HT-hTS, HT-hTS F59A, HT-hTS Y202A and (**B**) of the same mutants incubated with dUMP for 1 h on ice followed by circular dichroism spectroscopy. Melting temperature (T_m) values determined for the transitions are reported in the tables.

3.2.2 Structural characterization of hTS Y202A and hTS Y202A in complex with the dissociative inhibitor compound 1

The crystal structures of HT-hTS Y202A and its complex with the dissociative inhibitor compound 1 have been obtained to resolutions of 2.40 and 2.55 Å, respectively. Both crystals belong to the $P3_121$ space group with a single chain found in their asymmetric unit (ASU) (*Table 3.2*), analogously to the wt enzyme [58]. The HT-hTS Y202A mutant crystallizes under high salt conditions, yielding crystals isomorphous to those of HT-hTS. The initial 13 N-terminal residues of the hTS sequence (and the 12 residues belonging to the His^6-tag) are not visible in the crystal structures, coherently with former structural data on this enzyme [58, 65, 70]. hTS has a mixed α/β structures and the dimer interface is generated by two large, reversely symmetrical mixed-sheets [43, 122]. The alanine point-mutation does not affect the overall fold of the enzyme, which is conserved with respect to HT-hTS (PDB code 3N5G), as also suggested by CD spectra. The comparison of the structures results in a rmsd upon $C\alpha$ matching of 0.59 Å (maximal displacement of 3.87 Å on residue Gly136). The structure of the Y202A mutant clearly shows that the enzyme is in the inactive conformation, having the catalytic loop 181-197 shifted toward the dimer interface (*Figure 3.8*). This conformation of the enzyme is typically observed when the protein is crystallized under high salt conditions [59, 123]. The inspection of the Fourier difference maps clearly shows the mutation of Tyr202 to Als202 (*Figures 3.8 and 3.9 A*). Ala202 is H-bonded to Ser209 on the next β-strand, interactions mediated by their backbones and thus not affected by the mutation. Residue 202 is located at the periphery of hTS dimer, on the fourth strand of the interface β-sheet (*Figure 3.9 A*), where it faces Phe59' on the other protomer (*Figure 3.9 A*). In the structure of the mutant, Ala202 takes hydrophobic interaction with the facing Phe59' (*Figure 3.9 B*). On the other hand, in the wild type of enzyme, the side chain of Tyr202 is stacked to the aromatic moiety of the facing residues (*Figure 3.9 B*). Superimposition between our HT-hTS Y202A and that one previous deposited

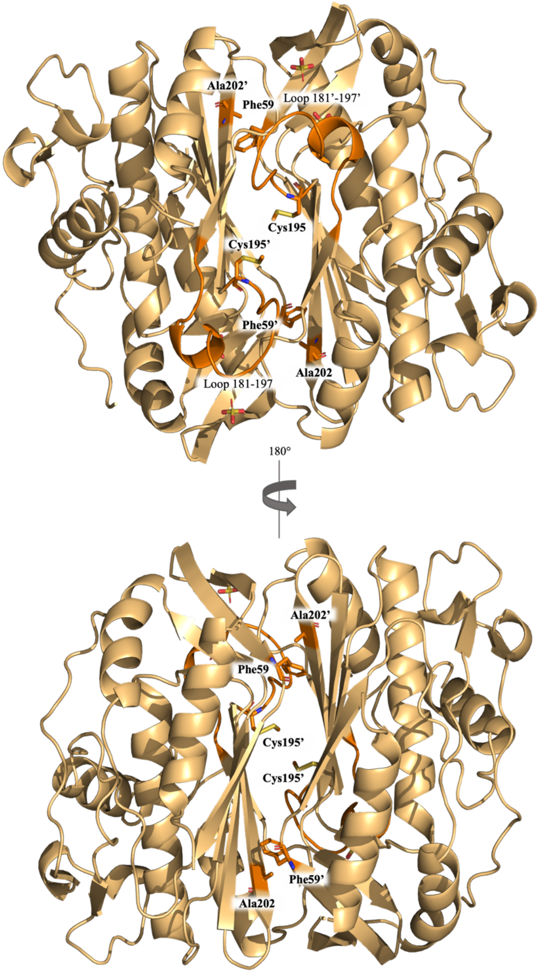


Figure 3.8: Cartoon representation of hTS Y202A enzyme homodimer (subunit A and B are colored in lightorange). both subunits assume the inactive conformation (loop 181-197 showed in orange), showing the catalytic Cys195 (in orange and carbon atom sticks) exposed towards hTS dimer interface. The catalytic Cys195 is modified as S,S-(2-hydroxyethyl)thiocysteine (SCH195, in orange and carbon atom sticks) in both subunits. The mutated residue Ala202 and the facing Phe59 of the other protomer are displayed in orange and carbon atom sticks. Sulfate anions are displayed in sticks in the Figure.

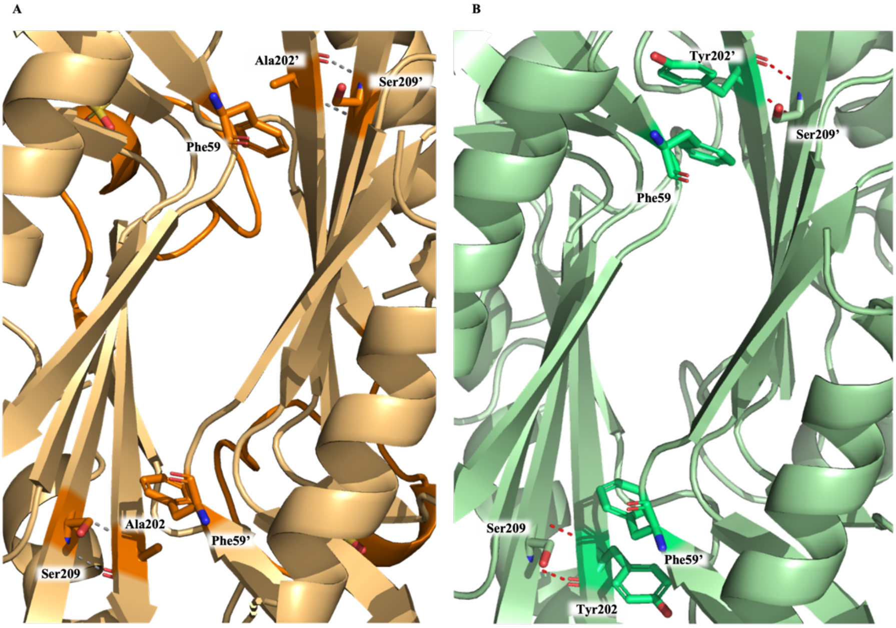


Figure 3.9: Interface view of hTS Y202A variant (light orange cartoon) (**A**) and wt hTS enzyme (pale green cartoon) (**B**). Both enzymes are represented in cartoon and residues in position 59 and 202 are displayed in orange and carbon atoms (**A**) or green and carbon atoms sticks (**B**). The mutated residue Ala202 (**A**), as well as Tyr202 (**B**) establishes two backbone-backbone H-bonds (grey (**A**) and red dashes (**B**)) with S209 on the next β-strand. (**A**)Ala202 residue takes hydrophobic interaction with the facing F59'. (**B**) The side chains of Tyr202 are stacked to aromatic moiety of the facing residues Phe59'.

for the same mutant (PDB id 4JEF) [70] displays that the X-ray structures are identical, as witnessed by the low value, 0.21 Å, of the rmsd upon Cα matching (maximal displacement of 3.12 Å, for residue Phe142). As expected, the point mutation Y202A is not enough itself to disrupt interface contacts in the hTS homodimer. In addition, in the structure of the HT-hTS Y202A we found two sulfate anions, respectively, deriving from the crystallization conditions. The first sulphate ion is salt-bridged to Arg50, Asn183, His208, Arg215, and Arg175'. Instead, the second anion is bound to Arg50, Arg78, Arg176, Arg185 and the backbone amine of Thr306. The positions and binding of these sulfate ions is identical to those formerly reported in the structures of wt-hTS (PDB id 3N5G) [58] and Y202A mutant (PDB id 4JEF) [70]. In this work we also tested a dimer disrupter inhibitor recently synthesized by the research group of Professor Maria Paola Costi (University of Modena and Reggio Emilia).

Compound 1 has IC_{50} values of 21.72 μM and 8.40 μM towards hTS wild type and hTS Y202A, respectively (unpublished data). Our crystallographic data show that compound 1 binds the periphery of the monomer-monomer interface of HT-hTS Y202A (*Figure 3.10*). The complex was determined by preparing the sample through the co-centrifugation method, whereas the presence of the inhibitor was not detected in the samples of the Y202A mutant prepared by the concentrated incubation method and in those of HT-hTS prepared by both methods. The comparison between the structures of HT-hTS Y202A and its complex with compound 1, shows that the binding of the inhibitor does not significantly alter the enzyme structure. This is supported by the low value, 0.45 Å, of the rmsd upon Cα matching (maximal displacement of 3.37 Å, for residue Gly136). The analysis of the interface shows that the monomer-monomer distances are not affected by the point mutation Y202A and by the presence of compound 1. In the Y202A structure, the inhibitor occupies an inverted-cone-shaped cleft, lined by residues 162-163, 166-167 and 176-178 of one protomer and the loop 189'-194' of the facing half, at the interface periphery (*Figure 3.10*). In this area, the inhibitor entails both hydrophobic and hydrophilic interactions with the surrounding residues inside the cleft. This binding region partially overlaps with the area buried by the LR peptide, the octapeptide formerly shown to act as hTS inhibitors by binding at the interface periphery [58].

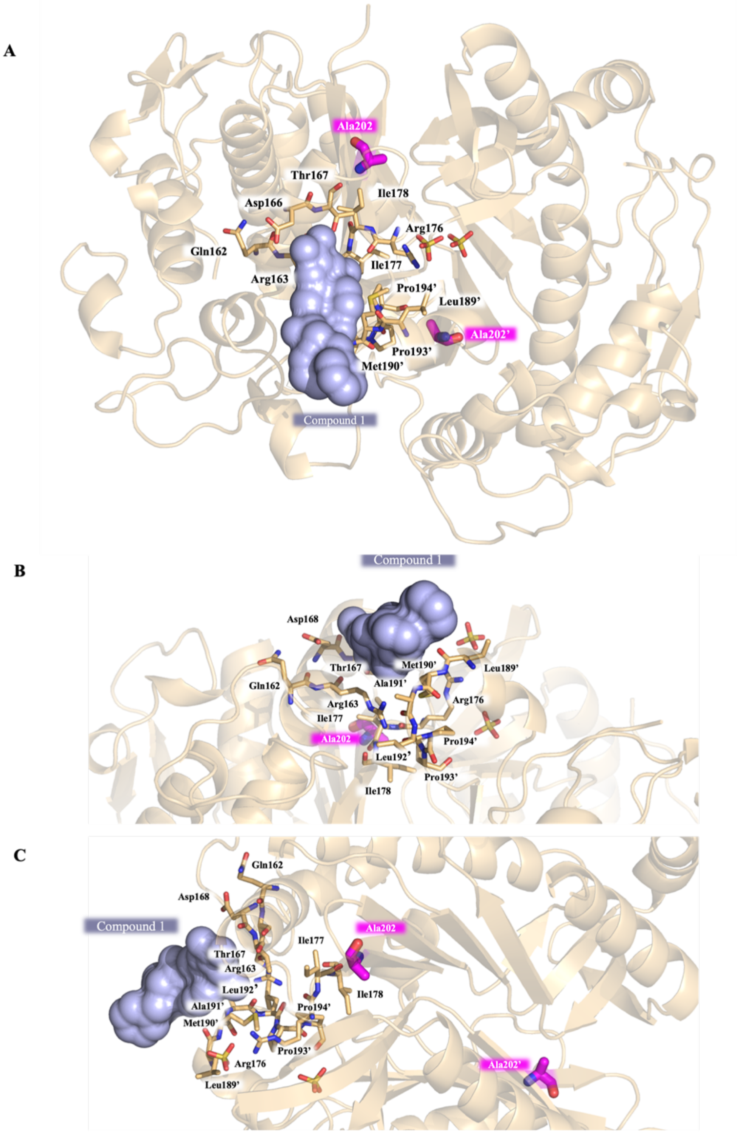


Figure 3.10: Three different views (**A**), (**B**) and (**C**) of the binding site of compound 1 (in light-blue surface) at the periphery of the dimer interface of the hTS Y202A variant (light orange cartoon). The mutated residue Ala202 is represented in sticks, magenta carbon atom. The inhibitor occupies an inverted-cone-shaped cleft (residues in sticks) at the interface periphery.

Chapter 4

Evidence of hTS dimer destabilization by interface mutations to charged amino acid residues

4.1 Materials and Methods

4.1.1 Cloning and Site-Directed Mutagenesis

The hTS charged variants Q62R, T251E and Q62R-T251E were generated by PCR-based site-directed mutagenesis, using partially overlapping primers [107] (*Table 4.2*; primers purchased from Sigma-Aldrich, Milan, Italy). The pQE80L-hTS plasmid [70] (*Figure 2.3*, see details in *Section* **??**) was used as template for the Q62R and T251E variant, whereas the hTS Q62R-T251E double mutant was generated using the pQE80L-hTS variant Q62R [108] plasmid as template. The Q62R mutation was obtained through PCR using the Expand High Fidelity polymerase (Roche Biochemicals, Basel, Switzerland), according to the published protocol[108], detailed in *Section 2.3*. For the other hTS mutants, the PCR reaction mixtures (50 μL) included 2 ng template DNA, 10 μM primer pair, 1 unit AccuPrime Taq DNA Polymerase High Fidelity and 1X AccuPrime PCR Buffer I, according to the manufacturer's protocol (Thermo Fisher Scientific). The reactions were performed through an initial denaturation step (94 °C, 30 sec), followed by 25 cycles of denaturation (94 °C, 30 sec), annealing (64 °C, 1 min) and extension (68 °C, 6 min). Afterwards, a 10-μL sample of each reaction mixture was incubated at 37 °C for 15 min with 1 μL of Fast Digest DpnI (Thermo Fisher Scientific), to allow digesting the methylated template DNA. The resulting samples were heat-shock transformed in chemically competent *E. coli* TOP10 cells and positive transformants were selected on LB-agar plates, supplemented with 100 mg L^{-1} ampicillin. For each variant, single colonies were cultured in LB medium added by 100 mg L^{-1} ampicillin overnight at 37 °C (220 rpm) and thus, used for plasmid extraction (E.Z.N.A.® DNA extraction Kit I, Omega Bio-Tek). The entire gene of each variant was sequenced to confirm the site-directed mutagenesis and to verify the absence of unwanted mutations (ATAC sequencing, Eurofins Genomics Italy).

Gene target	*Primer*	*Sequence 5'-3'*	*Product*
hTS WT	Fw Q62R	GCATGCGGGCGAGATATTCATTAC	*hTS Q62R*
	Rv Q62R	CTCGCCCGCATGCCGAAGACGC	
hTS WT	Fw T251E	ATCCACGAACTGGGTGACGCCC	*hTS T251E*
	Rv T251E	CCCAGTTCGTGGATAAAATCACCC	
hTS Q62R	Fw T251E	ATCCACGAACTGGGTGACGCCC	*hTS Q62R-T251E*
	Rv T251E	CCCAGTTCGTGGATAAAATCACCC	

Table 4.1: Sequences of forward (Fw) and reverse (Rv) primers used for mutagenic PCR to introduce the point mutations Q62R and T251E. The hTS variant Q62R and T251E were obtained using the pQE80L-hTS plasmid as template, whereas the double mutant Q62R-T251E was generated from the plasmid pQE80L-hTS variant Q62R. The mutated codons on both strands are in red.

4.1.2 Recombinant Protein Expression and Purification

His^6-tag hTS Q62R (HT-hTS Q62R), hTS T251E (HT-hTS T251E) and the hTS Q62R-T251E (HT-hTS Q62R-T251E) were expressed in the *E. coli* BL21(DE3) strain, according to the protocol published for the Q62R variant [108], detailed in *Section 2.3.* Briefly, bacterial cells were cultured in ZYP-5052 auto-induction medium [124] supplemented with 100 mg L^{-1} ampicillin, at 20 °C for 48 h. Cells, harvested by centrifugation, were resuspended in buffer A (30 mM NaCl and 50 mM HEPES, pH 7.5) containing 20 mM imidazole, 0.2 mM PMSF, 0.5 mg mL^{-1} lysozyme, and then disrupted by sonication after 1 h incubation on ice. The soluble cellular fraction, clarified by centrifugation (13500 g, 1 h, 8 °C), was purified by nickel affinity chromatography (HisTrap FF 5 mL column, GE Healthcare) using a three-step gradient elution protocol relying on imidazole as competitive agent (80 mM, 250 mM and 500 mM). Fractions containing the target proteins, identified by SDS-PGE, were pooled and extensively dialyzed in buffer A. The resulting samples were concentrated (Pierce^{TM} Protein Concentrator PES, molecular weight cut-off 10 kDa) and then purified by size exclusion chromatography on a HiLoad 16/600 Superdex 75pg column (GE Healthcare). The high purity of the target proteins was confirmed as >98% by SDS-PAGE and MALDI-TOF mass spectrometry (*vide infra*) analyses.

4.1.3 Mass Spectrometry analysis

Matrix Assisted Laser Desorption Ionization - Time of Flight (MALDI- TOF) mass spectrometry analyses were performed on all purified mutants. Each sample contained a 20 μM protein solution in ultrapure water. Sinapinic acid was used as matrix, in a 1:1 water/acetonitrile solution added by 0.1% trifloroacetic acid (TFA). Spectra were recorded on a MALDI TOF Ultraflex III TOF/TOF200 (Mass Spectrometry facility, Toscana Life Science).

4.1.4 Kinetic Activity Assay

Kinetic analyses on HT-hTS Q62R and HT-hTS T251E mutants were performed according to the protocol reported for wild-type hTS and alanine variants [70] by the research group of Prof Maria Paola Costi (University of Modena and Reggio Emilia Modena). The kinetic profile of the HT-hTS Q62R-T251E double variant was evaluated spectrophotometrically, using a UV-visible spectrophotometer (Lambda 900/Perkin Elmer Instrument), according to the same protocol [70, 125]. Briefly, reactions were performed in a 1-mL assay buffer mixture (50 mM TES, pH 7.4, 25 mM $MgCl_2$, 6.5 mM formic acid, 1 mM EDTA, 75 mM β-mercaptoethanol) and started by adding the substrate. To determine kinetic parameters, each reaction was monitored for 3 min following the increase in absorbance at 340 nm due to mTHF oxidation. K_M (Michaelis-Menten constant) values were determined for both dUMP and mTHF by varying their assay concentrations (in the ranges 3–180 μM and 5–80 μM, respectively), keeping fixed the enzyme concentration (2 μM). On the other hand, the k_{cat} was determined using a fixed concentration of dUMP and mTHF (120 μM and 55 μM, respectively) and varying the enzyme concentration (in the range 0.2-4 μM). Measurements were performed in triplicates.

4.1.5 Circular Dichroism (CD) thermal denaturation analysis

CD spectra of the variants HT-hTS Q62R, HT-hTS T251E, and HT-hTS Q62R-T251E were recorded in the wavelength range 250-200 nm to verify their folding in comparison to HT-hTS. Each sample consisted of a 200 μL protein solution at the concentration of 20 μM in 10 mM HEPES pH 7.5 and 100 mM NaCl. Spectra sere recorded using 0.1 cm quartz cuvette at room temperature (buffer signals were subtracted during spectra recording). For each variant we also determined the stability profile through thermal unfolding experiments by monitoring the far-UV CD signal at 220 nm during sample heating from 25 °C to 80 °C (2 °C min^{-1} rate). A second set of experiments was performed on each variant incubated with 1 mM dUMP for 30 min on ice. Measurements were performed in triplicates. Data were analyzed using the software GraphPad Prism 7 in non-linear regression using the Boltzmann sigmoidal function for T_m determination.

4.1.6 Crystallization

The recombinant HT-hTS Q62R and HT-hTS Q62R-T251E were concentrated to 20 mg mL^{-1}, while HT-hTS T251E to 15 mg mL^{-1} in 30 mM NaCl and 50 mM HEPES, pH 7.5 and stored at -20 °C until required. Crystallization screens (JBSB 1–4 and JBSC 6 from Jena Bioscience and PEG/Ion, Index, Crystal Screen 1 and 2, and Grid screen Ammonium Sulfate from Hampton Research) were performed on these protein variants using the vapor diffusion sitting drop technique at room temperature [109]. Drops, consisting of 2 μL protein (with or without 20 mM β-mercaptoethanol) and 2 μL precipitant solutions, were equilibrated against a 200 μL reservoir. For HT-hTS Q62R, crystal growth was observed in 10–14 days using the JBSC6 solution C4 (2.0 M ammonium sulfate, 100 mM TRIS pH 8.5) as precipitant. For HT-hTS T251E, crystal growth was observed in 10-15 days using the JBSB3 solution D2 (1.4 M tri-Sodium Citrate, 100 mM TRIS pH 7.5) as precipitant. On the other hand, crystals of HT-hTS Q62R-T251E were obtained after 20-25 days using the Index solution 29 (60% v/v Tacsimate^{TM} pH 7.0). All variants were crystallized only in presence of 20 mM β-mercaptoethanol, added to the sample solution. The optimization of the crystallization conditions was performed using the hanging drop vapor-diffusion method [109] at 20 °C, by varying both salt concentration and buffer. Crystals, suitable for diffraction experiments, were obtained from drops prepared by mixing equal volumes of protein and precipitant solutions (Q62R: 2.0–2.2 M ammonium sulfate, 100 mM bicine pH 9.0; T251E: 1.0 M tri-Sodium Citrate, 100 mM TRIS pH 7.5; Q62R-T251E: 60% v/v Tacsimate$^{\mathrm{TM}}$ pH 7.0), equilibrated over 800 μL reservoir. Crystal growth (*Figure 4.1*) was observed within 15 days for all variants. Crystals of HT-hTS Q62R-T251E in complex with either dUMP or FdUMP were obtained by co-crystallization in drops consisting of equal volumes of protein (above protein solution added by either 20 mM dUMP or 4 mM FdUMP) and precipitant (55% v/v Tacsimate$^{\mathrm{TM}}$ pH 7.0) solutions equilibrated over 200 μL reservoir. Prior to data collection, crystals were singularly transferred to the cryoprotectant solution (20% v/v ethylene glycol or glycerol and 80% v/v precipitant solution) and flash frozen in liquid nitrogen.

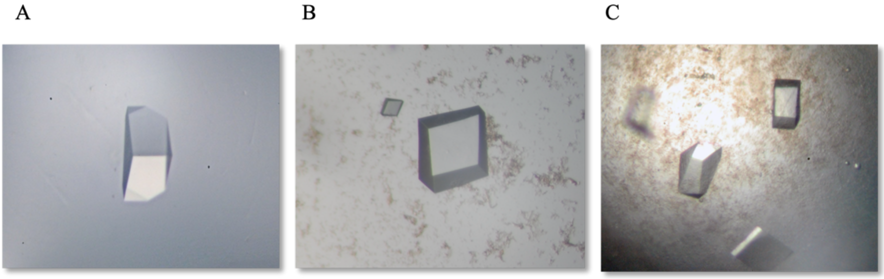


Figure 4.1: Crystals of HT-hTS Q62R (**A**), HT-hTS T251E (**B**) and HT-hTS Q62R-T251E (**C**).

4.1.7 Data collection, structure solution and refinement

X-ray crystallographic data were collected at 100 K using synchrotron radiation at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) beamline ID30B and at the Diamond Light Source (DLS, Didcot, UK) beamlines I02, I03 and I04. Reflection data were indexed and integrated using the program XDS [110] and scaled with SCALA [111, 112] from the CCP4 suite [113]. Data collection and reduction statistics are reported in *Tables 3.3 and 4.2*. Structures were solved by molecular replacement using the program Molrep [114] and one subunit of either hTS (active conformation; PDB code 1HVY

[126]) or HT-hTS Q62R (PDB code 6R2E) [108] as search model for the structures of the single and double variants, respectively. All structures were refined through the program REFMAC5 [115] from CCP4 suite [113]. The molecular graphic software Coot [116] was used for manual rebuilding and modelling of missing atoms in the electron density, and to add solvent molecules. Molecules of glycerol and ethylene glycol, and phosphate/sulfate anions (from production/crystallization conditions) were found in our structures and modelled according to the electron density. In all structures, the inspection of the Fourier difference map clearly evidenced the presence of a ligand inside the enzyme active site, identified as the folate derivative 5-formyl-tetrahydrofolate-di-L-glutamate (5-FTHF-di-L-Glu), according to former investigations on TS enzymes (*vide infra*). Furthermore, in the structures of HT-hTS Q62R-T251E in complex with either dUMP or FdUMP, we also observed additional peaks in the Fourier difference map, indicating their presence within the enzyme active site. The occupancy of all exogenous ligands was adjusted and refined to values resulting in atomic displacement parameters close to those of neighboring protein atoms in fully occupied sites. The final models were inspected manually, checked with Coot [116] and PROCHECK [117], and then validated by the PDB deposition tools. Data refinement statistics are reported in *Tables 4.3* and *4.4*. Structural figures were generated through the molecular graphic software PyMOL [118] and CCP4 mg [119].

4.1.8 Protein data bank (PDB) deposition

Coordinates and structure factors for HT-hTS Q62R were deposited in the Protein Data Bank under the accession code 6R2E [108].

	HT-hTS Q62R	**HT-hTS T251E**	**HT-hTS Q62R-T251E**
Diffraction source	ID30B (ESRF)	I02 (DLS)	I03 (DLS)
Wavelength (Å)	0.96861	0.97625	0.976230
Temperature (K)	100	100	100
Detector	Dectris Pilatus3 6M	Pilatus3 6M	Pilatus3 6M
Crystal-detector distance (mm)	460.0	368.389	287.349
Exposure time per image (s)	0.20	0.20	0.20
Space group	$P2_122_1$	$P2_122_1$	$P2_122_1$
No. of heterodimers in the ASU	8	8	8
a = b, c (Å)	139.94; 167.07; 189.97	142.00; 167.85; 190.30	140.09; 165.76; 190.12
Resolution range (Å)	94.99-2.55	95.15-2.70	29.60-2.25 (2.37-2.25)
Total no. of reflections	861677 (109995)	538835 (77852)	1726632(212452)
No. of unique reflections	145044 (20911)	122656 (17836)	207779 (28926)
Completeness (%)	99.9 (99.6)	98.3 (98.7)	99.3 (95.4)
Redundancy	5.9 (5.3)	4.4 (4.4)	8.3 (7.3)
$\langle I/\sigma(I)\rangle$	6.9 (2.1)	4.8 (2)	9.4 (2.7)
R_{meas}	0.178 (0.782)	0.154 (0.530)	0.110 (0.729)
Overall B factor from Wilson plot (Å^2)	26.83	34.2	41.83

Table 4.2: Data collection and processing statistics for HT-hTS Q62R, T251E and Q62R-T251E structures. Values in parentheses are for the highest resolution shell.

4.2 Results and Discussions

4.2.1 Expression and purification of HT-hTS wild type and interface variants

All HT-hTS interface variant were expressed as recombinant proteins having a non-cleavable N-terminal His^6-tag, resulting in whole sequences of 325 residues (sequence numbering refers to the wild-type hTS, UniProt KB id P04818) [58]. Despite the interface mutations, all variants were mainly expressed in the soluble cellular fraction of *E. coli* BL21(DE3) cells. Nonetheless, with respect to the production of HT-hTS, we had to apply slightly different bacterial culturing conditions[108] (lower temperature and longer incubation times). The purification protocol relied on a single-stage of nickel affinity chromatography (representative chromatogram and SDS-PAGE are shown in *Figure 4.1 A and B*), followed by SEC (representative chromatogram in *Figure 4.2*). The elution profile of all variants was consistent with the enzyme dimer assembly (*Figure 4.2*). Highly pure (>98%) protein samples were obtained by these procedures,

	HT-hTS Q62R-T251E	
	dUMP	**FdUMP**
Diffraction source	I04 (DLS)	I03 (DLS)
Wavelength (Å)	0.97950	0.976230
Temperature (K)	100	100
Detector	Pilatus 6M-F	Pilatus3 6M
Crystal-detector distance (mm)	252.778	270.125
Exposure time per image (s)	0.20	0.20
Space group	P21221	P21221
No. of heterodimers in the ASU	8	8
a = b, c (Å)	140.29,166.13,190.63	139.94,165.76,190.22
Resolution range (Å)	29.68-2.80 (2.95-2.80)	29.61-2.60 (2.74-2.60)
Total no. of reflections	848687 (122059)	1153883 (170786)
No. of unique reflections	109985 (15922)	136074 (19660)
Completeness (%)	99.9 (100.0)	99.9 (100.0)
Redundancy	7.7 (7.7)	8.5 (8.7)
$\langle I/\sigma(I)\rangle$	10.9 (2.9)	10.7 (3.3)
Rmeas	0.119 (0.631)	0.106 (0.617)
Overall B factor from Wilson plot (Å^2)	43.81	56.73
RSCC ligand		

Table 4.3: Data collection and processing statistics for the structures of HT-hTS Q62R-T251E variant in complex with dUMP or FdUMP. Values in parentheses are for the highest resolution shell.

	HT-hTS Q62R	**HT-hTS T251E**	**HT-hTS Q62R-T251E**
Resolution range (Å)	95.17-2.55 (2.62-2.55)	95.15-2.70 (2.85-2.70)	29.60-2.25 (2.37-2.25)
No. of reflections, working set	137694 (10032)	115974 (8536)	197210 (13252)
No. of reflections, test set	7182 (554)	6123 (428)	10436 (698)
R_{cryst}	0.196 (0.315)	0.231 (0.334)	0.233 (0.329)
R_{free}	0.258 (0.389)	0.277 (0.412)	0.283 (0.35)
No. of non-H atoms			
Protein	18550	18587	18476
Ion	93	40	40
Ligand (5-FTHF or 5-FTHF-di-Glu)	284	344	344
Water	2430	903	1089
Total	21357	19829	19331
R.m.s deviations			
Bonds (Å)	0.012	0.008	0.009
Angles (°)	2.167	1.746	1.792
Average B factors (Å2)	44.01	34.28	46.10
Estimate error on coordinates based on R value (Å)	0.326	0.466	0.221
Ramachandran plot			
Most favored (%)	96.5	96.1	95.6
Allowed (%)	3.5	3.9	4.4

Table 4.4: Refinement statistics for HT-hTS Q62R, T251E and Q62R-T251E structures. Values in parentheses are for the highest resolution shell.

	HT-hTS Q62R-T251E	
	dUMP	**FdUMP**
Resolution range (Å)	29.68-2.80 (2.95-2.80)	29.61-2.60 (2.74-2.60)
No. of reflections, working set	104505 (7660)	129127 (9422)
No. of reflections, test set	5425 (396)	6822 (0.299)
R_{cryst}	0.24986 (0.333)	0.21035 (0.299)
R_{free}	0.33761 (0.401)	0.27969 (0.369)
No. of non-H atoms		
Proteins	18978	18496
Ligands	120 (dUMP), 344 (5-FTHF-di-Glu)	168 (FdUMP), 344 (5-FTHF-di-Glu)
Water	1450	1474
Total	20892	20482
R.m.s. deviations		
Bonds (Å)	0.007	0.008
Angles (°)	1.663	1.825
Average B factors (Å^2)	45.013	50.057
Estimate error on coordinates based on R value (Å)	0.669	0.368
Ramachandran plot		
Most favored (%)	92.8	95
Allowed (%)	7.2	5

Table 4.5: Refinement statistics for HT-hTS Q62R-T251E variant in complex with dUMP or FdUMP.

as confirmed by SDS-PAGE analysis (*Figure 4.2 B*) and MALDI-TOF mass spectrometry (*Figure 4.2 A-C*; HT-hTS Q62R theoretical molecular weight MW = 37142.46 Da, determined MW = 37072.09 HT-hTS T251E theoretical MW = 37142.46 Da, determined MW = 37121.79 Da; HT-hTS Q62R-T251E theoretical MW = 37170.52 Da, determined MW = 37170.55 Da). The final production yields ranged from ≈50 mg L^{-1} to ≈100 mg L^{-1} bacterial culture (HT-hTS Q62R: ≈80 mg L^{-1}; HT-hTS T251E: ≈50 mg L^{-1}; HT-hTS Q62R-T251E: ≈100 mg L^{-1}).

4.2.2 Protein characterization: thermal stability and kinetic analyses

Former investigation on hTS have shown that residues placed on the monomer-monomer interface are crucial for the homodimer assembly of this enzyme [43, 70]. Here, we have generated three novel hTS interface variants in which charged residues (Arg and/or Glu) have been placed in critical positions of the hTS homodimer interface. Notably, the mutated amino acids at positions 62 and 251 face themselves on the cognate subunit. To assess the influence of the introduced interface mutation on the enzyme stability and catalytic efficiency, all variants were subjected to CD thermal denaturation and kinetic analyses in comparison to HT-hTS. At variance with the thermal denaturation profile of HT-hTS that showed a biphasic transition (*Figure 4.5 A, curve in green*), monophasic transitions were recorded for the interface variants (*Figure 4.4 A*). The introduced interface mutations induce reductions of the T_m values, resulting of 51.26 (± 0.08) °C [108] and 49.87 (± 0.08) °C for the single variant Q62R and T251E, respectively, and of 51,4 (± 0,06) °C for the double variant Q62R-T251E (*Figure 4.4 A*), with respect to HT-hTS, having a T_m of 53.27 (± 0.09) °C (*Section 3.2.2*).

As formerly discussed in *Section 3.2.2*, after the addition of dUMP, HT-hTS undergoes to a stabilization effect, leading to an increment of ≈3.5 °C in its T_m value, that reaches 56.77 (± 0.08) °C. Surprisingly, the T_m values of all interface variance are unaltered upon substrate addition, being 51.11 (± 0.10)

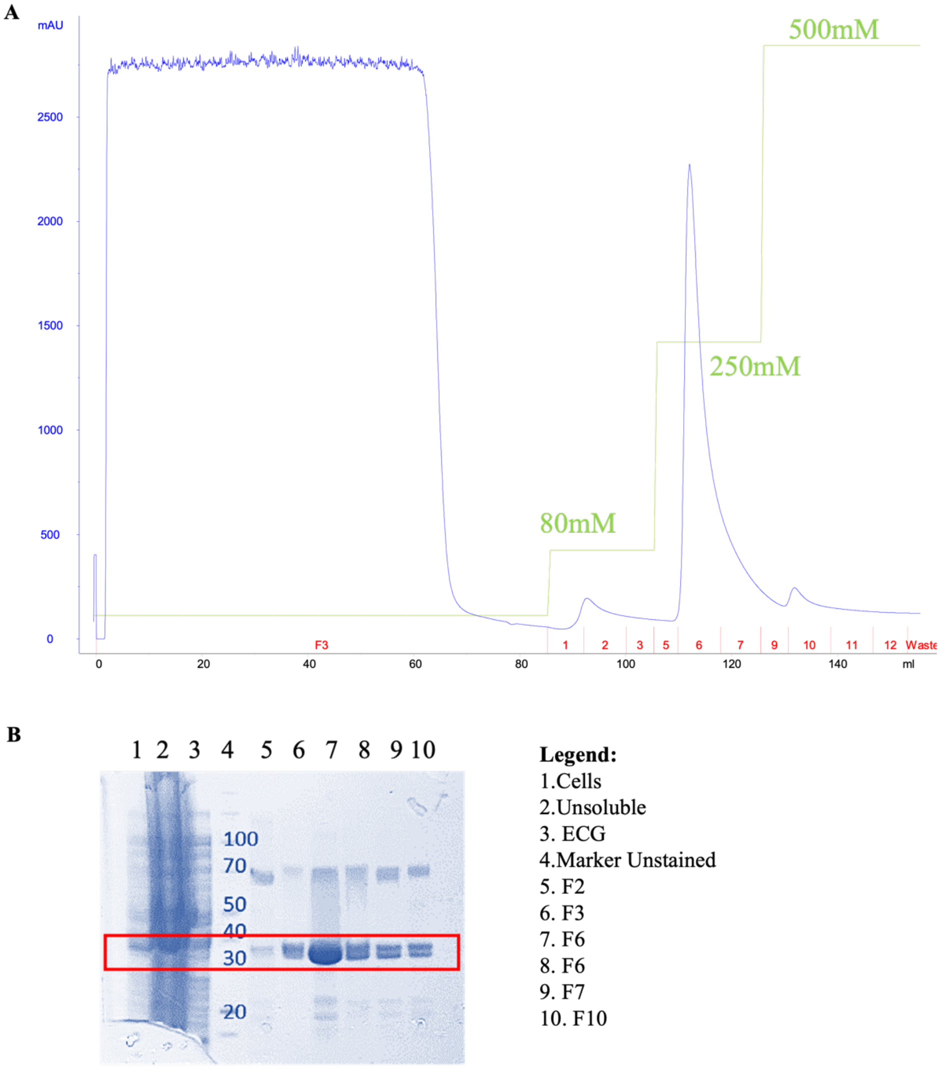


Figure 4.2: (**A**) Representative chromatogram of the nickel- affinity chromatography (HisTrap FF 5mL column, GE-Healthcare) performed to purify HT-hTS Q62R-T251E. The UV_{280nm} monitoring and imidazole concentration are shown as blue and green lines, respectively; the elution fractions are indicated on the x-axis as red marks. (**B**) SDS-PAGE analysis of the fractions collected from affinity chromatography (the protein ladder used for MW estimation is in lane 4).

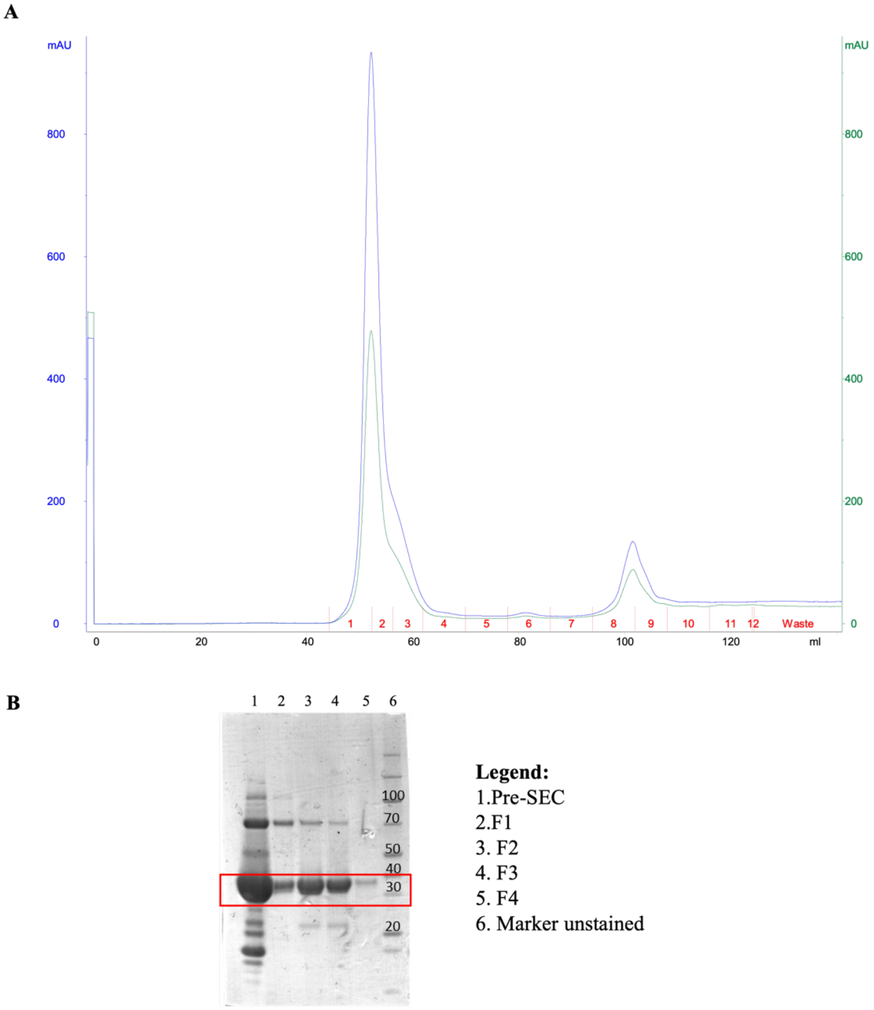
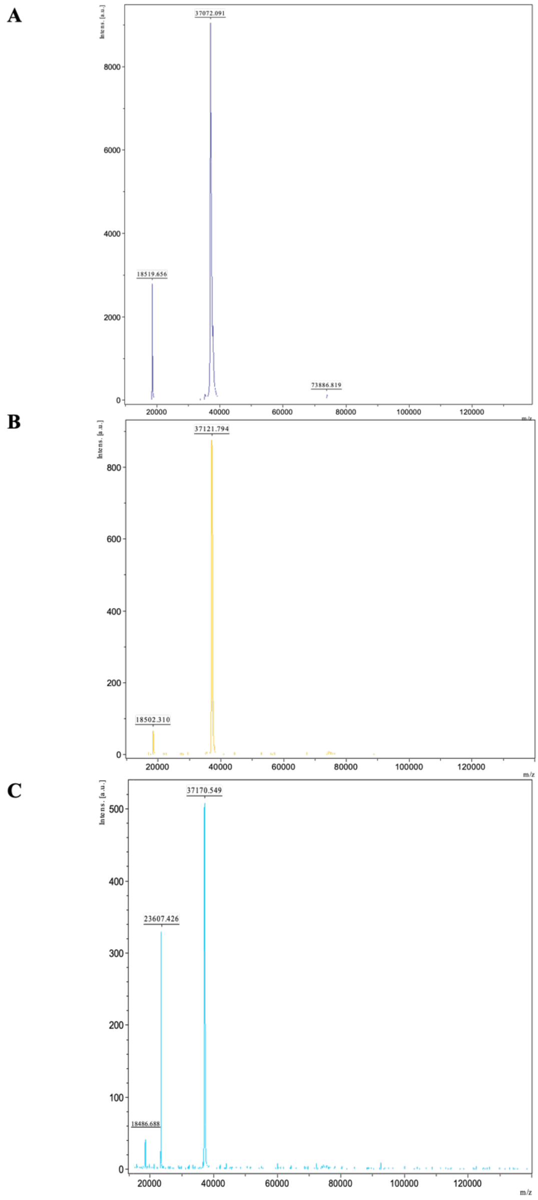


Figure 4.3: (**A**) Chromatogram of the HiLoad 16/600 Superdex 75pg column (GE-Healthcare) purification step. The UV_{280nm} and UV_{260nm} monitoring of protein elution is in blue and green respectively. Factions are indicated by red marks on the x-axis. (**B**) SDS-PAGE analysis of the fractions collected from SEC. The protein ladder is in lane 6.



m/z	$MW_{theoretical}$
18519.656	18571.230 (½ MW)
37072.091	37142.460 (MW)
73886.819	74284.920 (2 MW)

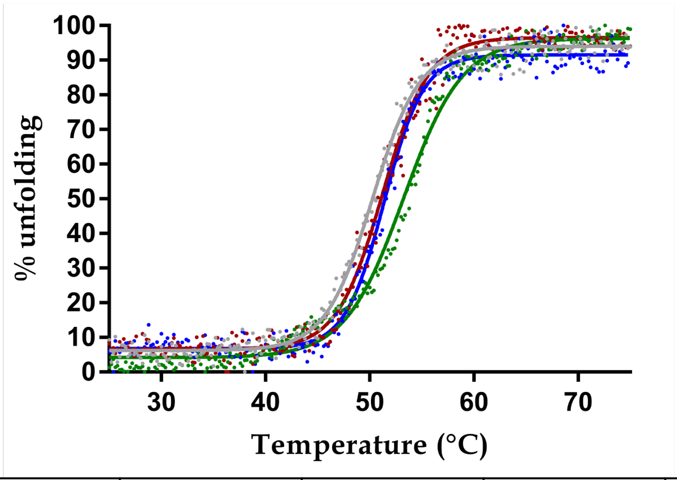
m/z	$MW_{theoretical}$
18502.310	18571.230 (½ MW)
37121.794	37142.460 (MW)

m/z	$MW_{theoretical}$
18486.688	18585.275 (½ MW)
37170.549	37170.550 (MW)

Figure 4.4: MALDI-TOF mass spectra of HT-hTS Q62R (**A**), HT-hTS T251E (**B**) and HT-hTS-Q62R-T251E (**C**).

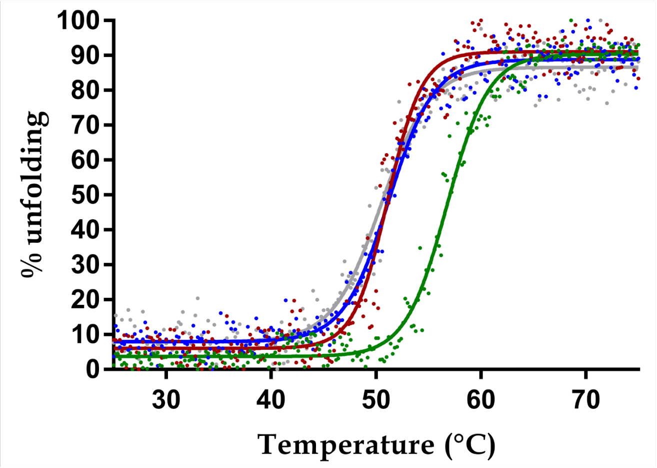
°C and 50.51 (± 0.15) °C for the single variant Q62R and T251E, respectively, and 51.22 (± 0.09) °C for the double variant Q62R-T251E (*Figure 4.4 B*). The comparison of the T_m values recorded in presence of dUMP, shows reduction of ≈ 5-6 °C for all interface variants with respect to HT-hTS (*Figure 4.4 B*). From these results we can deduce that the interface variants are not subjected to the stabilization effect induced by the substrate dUMP, attributed to the switch of the protein to the active conformation.

As expected, the point mutations introduced at the hTS interface deeply affected also the enzyme catalytic activity. To evaluate these effects, the kinetic properties of the mutants have been compared with the kinetic parameters formerly reported for HT-hTS [70] (K_m dUMP: 10 ± 1 μM; K_m mTHF: 6 ± 1 μM; k_{cat}: 1.00 ± 0.01 s^{-1}; summarized in *Table 4.6*). The affinity for the substrate is almost unaffected by the mutations, as indicated by the analogous K_M values for dUMP (*Table 4.6*). On the other hand, a slight alteration is observed for cofactor affinity, resulting reduced by ≈ 3-3.5 times in the interface mutants with respect to HT-hTS (K_m of 16 ± 1 μM and 19 ±1 μM for the single variant Q62R and T251E, respectively, and of 21 ±1 μM for the double variant Q62R-T251E; *Table 4.6*). Furthermore, all variants displayed a 20-30 times decreased turnover rate with respect to HT-hTS [70] (k_{cat} of 0.16 ± 0.06 s^{-1} and 0.04 ± 0.03 s^{-1}, for Q62R and T251E, respectively, and of 0.03 ± 0.04 s^{-1} for Q62R-T251E; *Table 4.6*). This reflects also on the catalytic efficiency, leading to significantly lower k_{cat}/K_m values for all interface variants (see *Table 4.6*).



Protein sample	HT-hTS	HT-hTS Q62R	HT-hTS T251E	HT-hTS Q62R-T251E
Melting Temperature (Tm)°C	53.27 ± 0.09 °C	51.26 ± 0.08 °C	49.80 ± 0.08 °C	51.4 ± 0.06 °C

B



Protein sample	HT-hTS + dUMP	HT-hTS Q62R+ dUMP	HT-hTS T251E+ dUMP	HT-hTS Q62R-T251E + dUMP
Melting Temperature (Tm)°C	56.77 ± .,08 °C	51.11 ± 0.10 °C	50.20 ± 0.08 °C	51.2 ± 0.09 °C

Figure 4.5: Thermal unfolding transition curves of HT-hTS, HT-hTS Q62R, HT-hTS Q62R-T251E (**A**) and of the same mutants incubated with dUMP (**B**) for 1 h on ice, followed by circular dichroism spectroscopy. Melting temperature (T_m) values determined for the transitions of each mutant in the absence (**A**) or in the presence (**B**) of ligand dUMP are reported in the tables.

	K_m dUMP (μM)	K_m mTHF (μM)	K_{cat} (s^{-1})	K_{cat}/K_m (dUMP) (M^{-1} s^{-1})	K_{cat}/K_m (mTHF) (M^{-1} s^{-1})
HT-hTS	6 ± 1	10 ± 1	1.00 ± 0.01	1.00×10^{-7}	1.67×10^{-7}
HT-hTS Q62R	8 ± 1	16 ± 1	0.16 ± 0.06	0.20×10^{-7}	0.10×10^{-7}
HT-hTS T251E	10 ± 1	19 ± 1	0.04 ± 0.03	0.4×10^{-8}	0.2×10^{-8}
HT-hTS Q62R-T251E	11 ± 1	21 ± 1	0.03 ± 0.04	0.3×10^{-8}	0.1×10^{-8}

Table 4.6: Kinetic characterization of His^6-tag human thymidylate synthase (HT-hTS) and of its interface variants HT-hTS T251E and HT-hTS Q62R-T251E. The turnover rate (K_{cat}), the K_m for dUMP, as well as that for mTHF were determined. Also, the value indicating the catalytic efficacy (k_{cat}/K_m) is reported in the table.

4.2.3 Structural characterization of the HT-hTS variants Q62R, T251E, and Q62R-T251E

4.2.3.1 Overall fold of the interface mutants Q62R, T251E, and Q62R-T251E

The structures of the interface variants Q62R (PDB id 6R2E) [108], T251E, and Q62R-T251E have been obtained at resolution ranging from 2.25 Å to 2.70 Å (*Table 4.2*), showing the dimeric quaternary assembly, typical of the wild-type enzyme (*Figure 4.6*). All crystals belonged to the primitive orthorhombic space group $P2_122_1$ including eight enzyme subunits (four enzyme dimers) in the ASU. In all structures, the four dimers (A-B, C-D, E-F, G-H) populating the ASU are nearly identical. The polypeptide chain was fully traced in all models, except for the starting 12 N-terminal residues belonging to the non-removable His_6-tag and the following 25 residues of the hTS sequence, disordered in all hTS structures reported so far [65]. From the initial phases of structure solution and refinement, it was evident that all variants were in the active conformation, with the catalytic loop 181-197 and the catalytic Cys195 exposed in the active site. The shape of the electron density in correspondence to the mutated residues at positions 62 and/or 251 proved the effective replacement of the amino acids at these sites. The side chains of the mutated residues were completely rebuilt in all enzyme subunits, according to the electron density. The structural comparison among all models shows that the overall fold is retained in all variants (rmsd of 0.17-0.37 Å upon Cα matching).

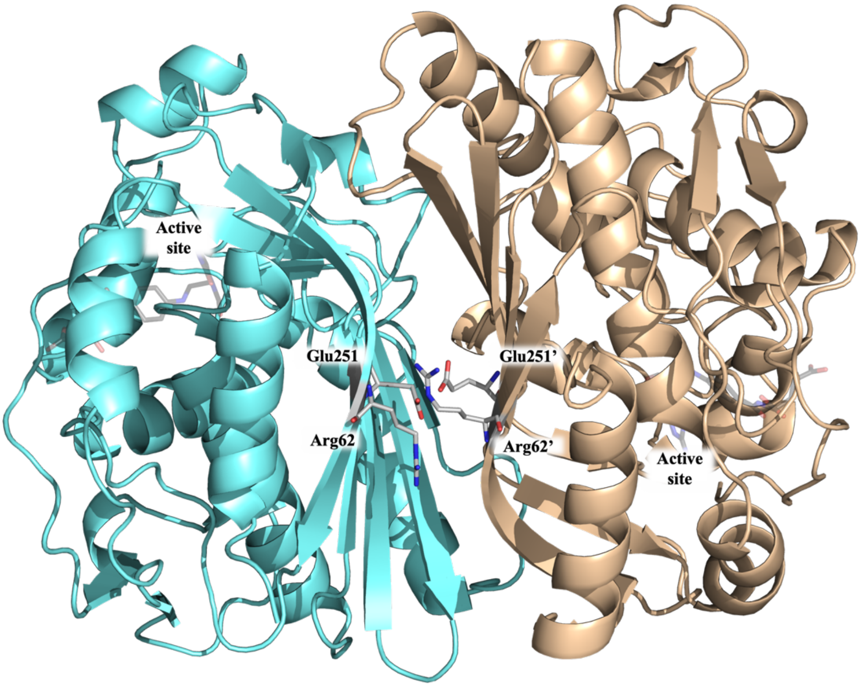


Figure 4.6: Cartoon representation of an enzyme homodimer (subunit A and B are colored cyan and wheat, respectively) of HT-hTS Q62R-T251E. The mutated residues Arg62 and Glu251 are displayed in sticks (gray and carbon atoms).

4.2.3.2 The dimer interface in HT-hTS Q62R, HT-hTS T251E and HT-hTS Q62R-T251E

Residues at positions 62 and 251 were selected to introduce charged interface mutations since they face themselves on the cognate subunit, thus possibly maximizing the perturbation effects induced by these modifications. In the structure of the Q62R mutant, the side chains of the two facing Arg62 point in two opposite directions at the dimer interface. This arrangement minimizes the electrostatic repulsions induced by the facing positive charges on the enzyme halves. In these orientations, the side chains of both arginines interact with the backbone carbonyl of Gly60 on the partner subunit [108]. Furthermore, on the two dimer halves, Arg62 is directed towards Arg64 of the same subunit, creating a positively charged area in which a sulfate anion (from the crystallization solution) is bound. In the structure of the HT-hTS Q62R-T251E double variant the same interaction with Gly60 is also visible, whereas the sulphate anion is not bound in this region (*Figure 4.8 B*). In the structures of the mutants Q62R [108] and Q62R-T251E, the distances measured between the Cα atoms of facing Arg62 are in the ranges 8.05 – 8.20 (± 0.46) Å and 8.16 – 8.39 (± 0.31) Å, respectively, in the four enzyme dimers found in the ASU. These distances are slightly increased with respect to those observed in the wild-type enzyme, where the facing Gln62 are placed $\approx$ 6.6 Å apart. The local shifts by $\approx$ 1.5 Å and $\approx$ 1.7 Å are spread over a large interface area, inducing a weakening of the dimer quaternary assembly of both Q62R single and double variants. As for the Q62R, also the T251E mutation determines mutual electrostatic repulsions of the two facing Glu251 and Glu251' residues; their negatively charged side chains point in the opposite directions in the structures of both HT-hTS T251E and HT-hTS Q62R-T251E. In each protomer, the side chain of Glu251 forms a H-bond with Gln211 (*Figure 4.8 A and B*).

Furthermore, in the structure of the double variant Q62R-T251E, the side chains of the mutated residues interact with each other, forming an intra-subunit interaction between the Arg62 guanidinium group the Glu251 carboxylate (*Figure 4.8 B*). In the structures of the mutants T251E and Q62R-T251E, the distances between the Cα atoms of facing Glu251 are in the range 7.90 - 8.10 (± 0.66) and 7.93 - 8.10 (± 0.31) Å, respectively, in the four homodimers presented in the ASU. These distances are analogous to those measured in wild-type hTS (PDB id 5X5D [84]), where Cα of facing Thr251 are 7.70 Å apart.

The structural data obtained on the three interface variants correlate with the destabilization effect reported by CD thermal denaturation analyses (*Section 4.2.2*). The point-mutations introduced at the hTS interfaces induce a slight aperture of the enzyme dimer, particularly evident in the areas surrounding Arg62 and Glu251, weakening the dimer quaternary assembly of the variants (*Figure 4.9*) and leading to a destabilization effect accounting for T_m reductions. The analysis of the interface areas in the mutant structures, performed through the PISA webserver (`http://www.ebi.ac.uk/pdbe/prot_int/pistart.html`), also supports a reduced stability of their dimer assembly [127]. Indeed, for the structure of

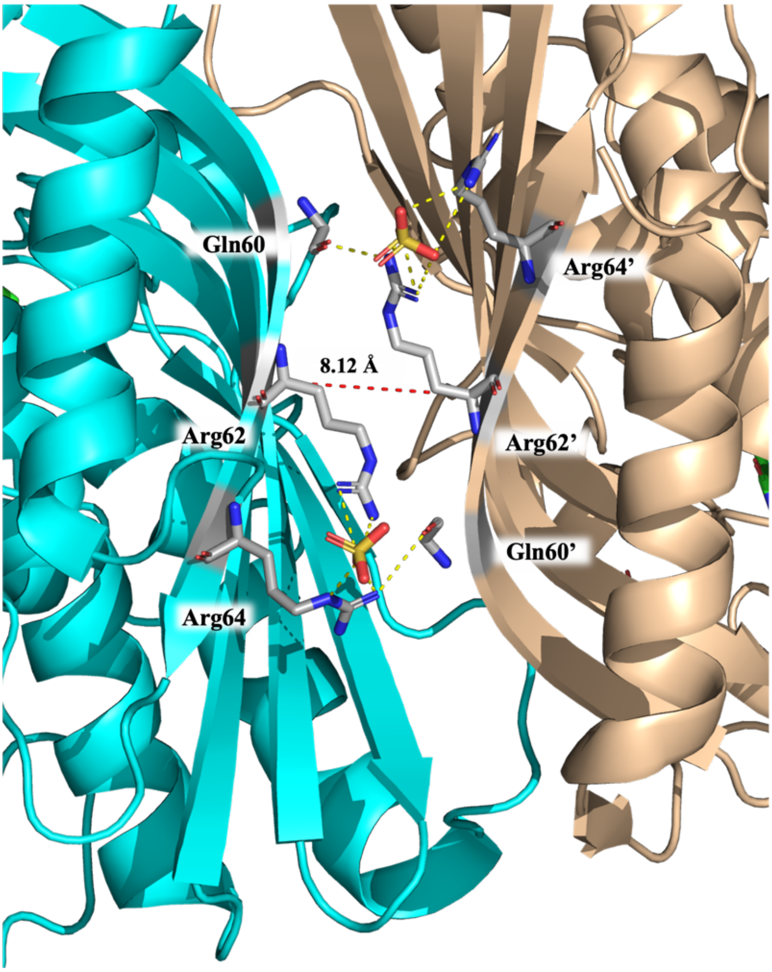


Figure 4.7: Interface view of HT-hTS Q62R variant (in cartoon, interacting residues in sticks; subunit A and B are colored cyan and wheat, respectively). Sulfate anions are displayed in sticks carbon atoms. The mutated Arg62 (grey sticks and carbon atoms) entails a H-bond (yellow dashes) with carboxylate of Gln60 in the same protomer; Arg62 is involved in some salt-bridge (yellow dashes) interaction with sulfate anions and Arg64 belonging to the same protomers. The mean distance between Cα atoms of facing Arg62-Arg62'(red dash) is 8.12 Å.

wild-type enzyme (PDB id 5X5Q [84]), it is calculated an average interface area of 2120.8 Å^2 and a Δ^{i}G (indicates the solvation free energy gain upon formation of the interface) of – 20.4 kcal mol^{-1}. On the other hand, slightly reduced interface areas are calculated for all variants, resulting of 2032.0 Å^2 and 1930.7 Å^2 for the Q62R [108] and T251E variants, respectively, and of 1954.4 Å^2 for the double mutant. This correlates with the increased Δ^{i}G values estimated to −18.7 kcal mol^{-1} and −18.8 kcal mol^{-1} for the Q62R [108] and

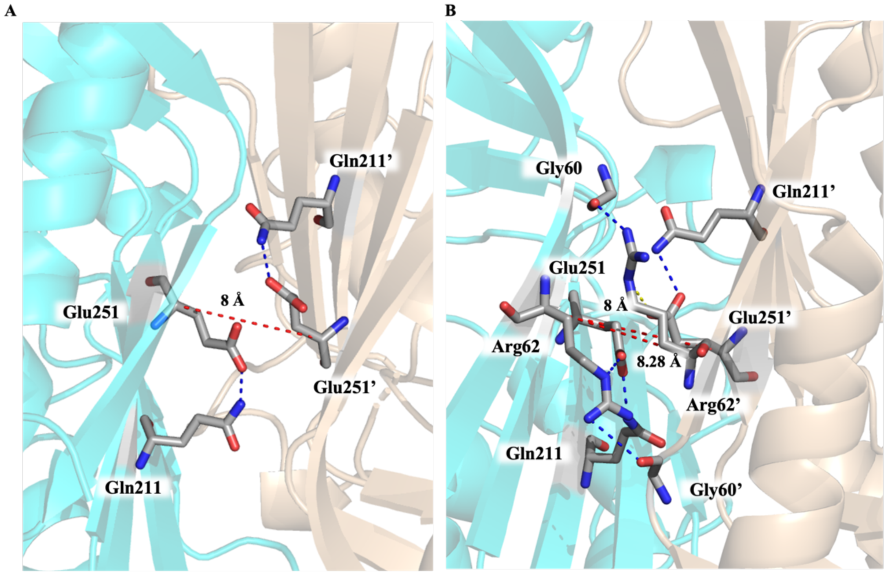


Figure 4.8: Interface view of HT-hTS T251E (**A**) and HT-hTS Q62R-T251E (**B**) variants (in cartoon, interacting residues in sticks; subunit A and B are colored cyan and wheat, respectively). The mutated T251E (grey sticks and carbon atoms) entails a H-bond (blue dashes) with carboxylate of Gln211 in the same protomer (**A**) and (**B**). (**B**) Arg62 is H-bonded (blue dashes) to Gly60 of the partner subunit and to Glu251 belonging to the same protomer. The mean distance between Cα atoms of facing Glu251E-Glu251E' (red dash) in the structure of T251E is 8 Å (**A**); the distances between the two facing Arg62-Arg62' and Glu251E-Glu251E'(red dash) are 8.28 and 8.05 Å, respectively.

T251E variants, respectively, and -17.6 kcal mol^{-1} for the double variant.

4.2.3.3 The active site in HT-hTS interface charged variants

The structural characterization of the three interface variants Q62R, T251E and Q62R-T251E revealed that are all in the active conformation. The analysis of the active site reveals the presence of a continuous electron density extending from the thiol of the catalytic Cys195, indicating its modification as S,S-(2-hydroxyethyl)thiocysteine (CME195) (Figure 23). This modification is due to the reaction with the β-mercaptoethanol added to the proteins prior to the crystallization experiments. The presence of this reducing agent is critical for hTS crystallization, indeed no crystal growth is observed without β-mercaptoethanol. Nearby CME195, a sulfate anion is observed, anchored to the guanidinium moieties of the four arginine residues Arg50, Arg215, Arg175',

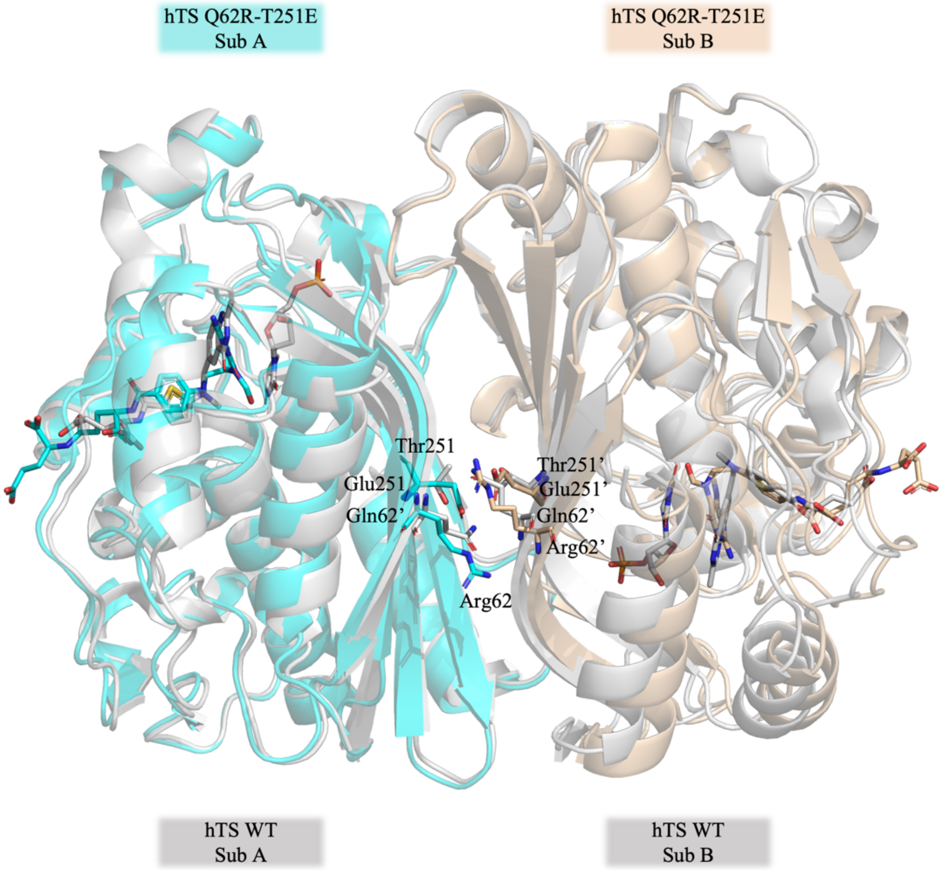


Figure 4.9: Structural comparison between the homodimer of HT-hTS Q62R-T251E (subunit A and B are colored cyan and wheat, respectively) and the wt enzyme (in grey; PDB id 1HVY [126]). The superimposition (performed by matching subunit A of both models) indicates a shift of the cognate subunit in the structure of the Q62R-T251E variant, with a slight opening of the quaternary dimeric structure.

and Arg176' (*Figure 4.10*). The positively charged pocket created by these arginines is responsible for recognition and binding of the dUMP phosphate moiety [84, 126], mimicked here by the sulfate anion.

In all structure, the active site is also further populated by an exogenous ligand, occupying the cofactor binding site. The shape of the electron density observed at this site is highly compatible with a folate-like molecule, and more precisely, with the reduced form for the cofactor having a bent conformation of the pyrazine ring of the folate pteridine. Furthermore, according to the observed electron density, this folate-like molecule has a biatomic substituent on

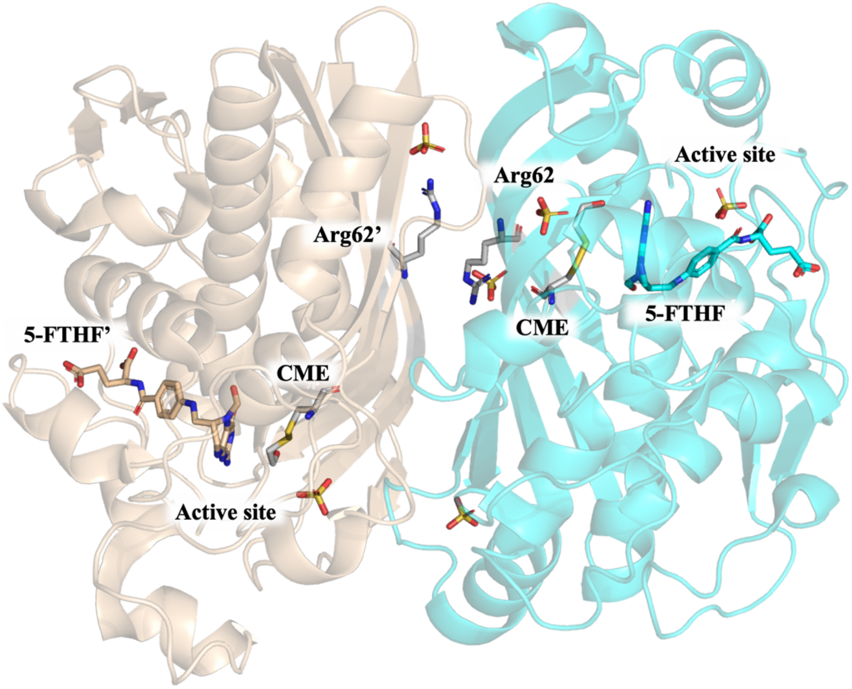


Figure 4.10: Cartoon representation of an enzyme homodimer (subunit A and B are colored cyan and wheat, respectively) of HT-hTS Q62R. The mutated residues Arg62 is displayed in sticks (gray and carbon atoms). The monomers adopt the active conformation showing the catalytic Cys195 exposed inside the catalytic cavity. The catalytic Cys195 is modified as S,S-(2-hydroxyethyl)thiocysteine (CME195, in grey and carbon atoms sticks) in all subunits. The cofactor analogue 5-formyl-6-tetrahydrofolate (5-FTHF, in sticks carbon atoms in wheat or cyan) is entrapped in the active site of both enzyme subunits.

the pteridine C5, consistent with either a hydroxymethyl moiety or a formyl group (5-ethyl derivatives of the cofactor are not known). The ligand was refined either as 5-hydroxymethyl-6-tetrahydrofolate (5-HMTHF) or 5-FTHF, without significant changes in the refinement quality indicators and in the resulting Fourier maps [108]. Moreover, the resolution of the structures (2.25 - 2.70 Å) did not allow us to discriminate between single and double C-O bonds, preventing further speculations on which THF-derivative is observed in this site. Nonetheless, we opted for 5-FTHF because this molecule is naturally formed inside cells [128] and its complexes with another hTS variant (PDB id 6QYQ) [129] and with the bacterial *Enterococcus faecalis* TS (*Ef*TS, PDB id

3UWL) were formerly reported [130]. Thus, we report the characterization of the HT-hTS Q62R in complex with the cofactor analogue 5-FTHF, copurified with the enzyme. In the structures of both HT-hTS T251E and HT-hTS Q62R-T251E, a further additional density extends to terminal glutamate moiety of 5-FTHF, suggesting that we are observing here its di-glutamate derivative, 5-formyl-6-tetrahydrofolate-di-L-glutamate (5-FTHF-di-Glu). After their synthesis, folate molecules are polyglutamylated, a modification that aids retention and compartmentalization in cells. *De facto* folate, as well as, folate-like molecules could have a γ-linked polyglutamyl tail of up to eight residues attached to the first glutamate [131]. Thus, both mono- and di-glutamate forms of 5-FTHF exist in cells and could be copurified with the TS variants. In the interface variant structure, both 5-FTHF and 5-FTHF-di-Glu occupy the cofactor binding site establishing a tight network of conserved H-bonds and van der Waals interactions with several residues exposed inside the catalytic cavity (*Figure 4.11*). The nitrogen N1 of the reduced pteridine is H-bonded to Asn112, whereas the nitrogen N2 establishes a water-mediated interaction with the same Asn112. The amine group forms a H-bond with Ala312 backbone carbonyl and Asp218 carboxylate. The latter residue also accepts a H-bond from N3 of the pteridine ring. Water-mediated interactions connect the ketone moiety of the reduced pteridine ring with the backbone nitrogens of Asp218 and Gly222, and the sidechains of Asn226. The formyl group on the pteridine N5 makes water-mediated interactions with Glu87 and His196 and a direct H-bond with the amide nitrogen of Asn226 (*Figure 4.11*). The p-aminobenzoate (pABA) portion of the cofactor analogues form a T-shaped interaction with Phe225. Moreover, the folate derivates make van der Waals interactions with Met311, Ile108, and Leu221 (*Figure 4.11*). The shared glutamate moiety of 5-FTHF and 5-FTHF-di-Glu is stabilized by a network of water-mediated interactions with the surrounding residues at the entrance of the active site, whereas the second glutamate moiety of 5-FTHF-di-Glu is mostly solvent exposed and establishes hydrophobic interactions with Phe80 (*Figure 4.11*).

The binding mode reported here for 5-FTHF is also the same observed in the

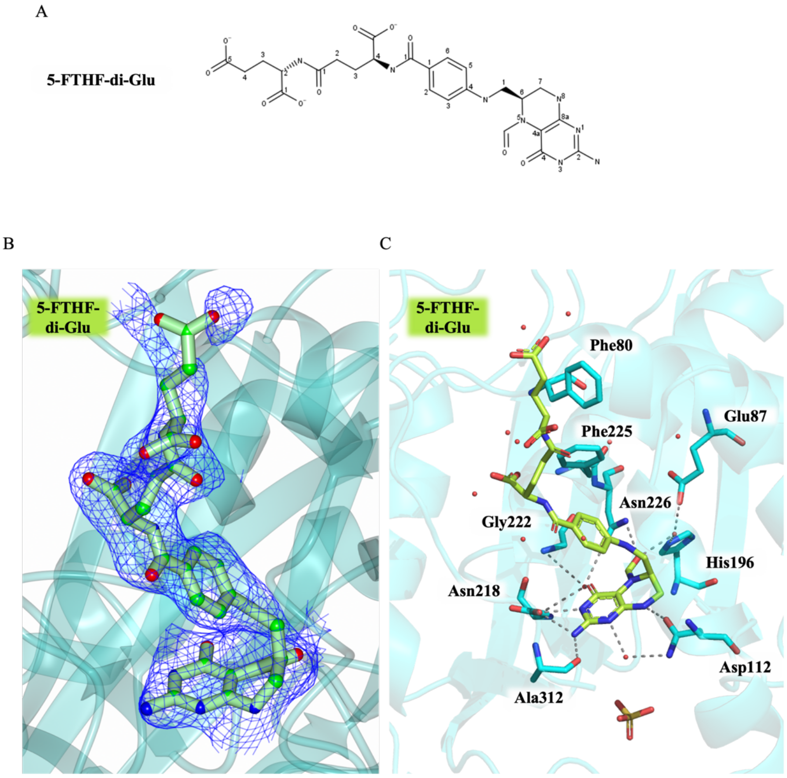


Figure 4.11: (**A**) Chemical structure of 5-formyl-tetrahydrofolate-di-glutamate (5-FTHF-di-Glu). (**B**) Cartoon representation of active site view of HT-hTS Q62R-T251E (in cartoon, interacting residues in cyan and carbon atoms sticks). The cofactor analogue 5-FTHF-di-Glu (in sticks, cyan carbon atoms) is entrapped inside the catalytic cavity by a tight network and it is surrounded by the omit map contoured at the 2.5 σ level. 5-FTHF is observed in all active sites of the four dimers found in the ASU. (**C**) Cartoon representation of active site view of HT-hTS Q62R-T251E (in cartoon, interacting residues in cyan carbon atoms sticks). The cofactor analogue 5-FTHF-diGlu establishes a tight network of H-bonds (grey dashed lines) and van der Waals interactions (water molecules in red spheres and water mediated interactions in grey dashed lines). Sulfate anions is displayed in sticks.

structures of the complexes with the hTS variant R175C (HT-hTS R175C; PDB id 6QYQ [129]) and with the bacterial enzyme *Ef*TS (PDB id 3UWL [130]), showing also retained interactions within the enzyme catalytic cavity. In all structures, the folate-like molecules populate the cofactor binding site despite the lack of substrate bound to its pocket (binary complexes). The comparison of our structural models with those of the ternary complexes hTS-dUMP-raltitrexed (PDB id 1HVY [126]) and hTS-FdUMP-raltitrexed (PDB id 6ZXO [58]) shows a different arrangement of 5-FTHF/5-FTHF di-L-glutamate and raltitrexed within the cofactor site (*Figure 4.12*). The reduced pteridine moiety of 5-FTHF/5-FTHF-di-L-glutamate is rotated by ≈30°with respect to the bicyclic system of raltitrexed, protruding in the substrate site (uracil pocket), occupied by dUMP/FdUMP in the ternary complexes [61, 84, 126, 132].

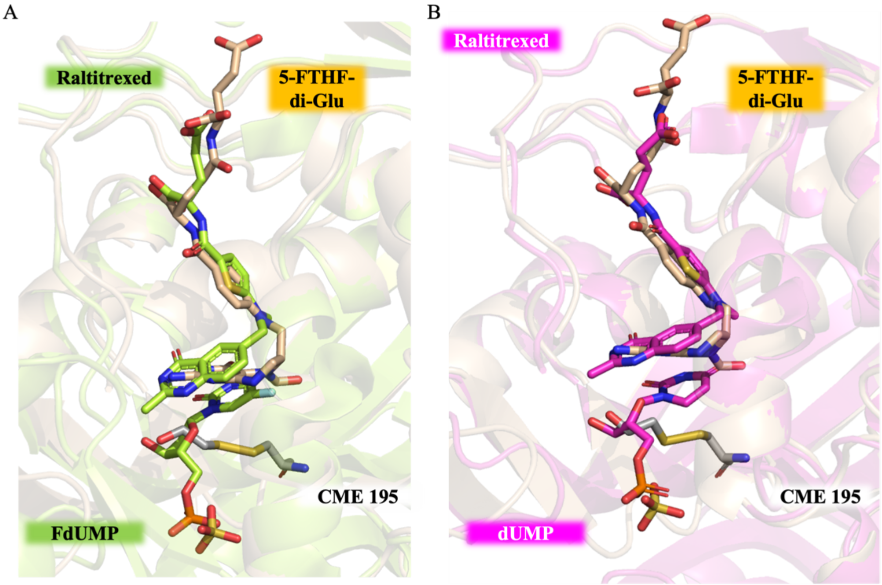


Figure 4.12: Active site view of the superimposition between the structures of HT-hTS T251E (wheat cartoon, sulfate anion in sticks) in complex with 5-FTHF-di-Glu (wheat carbon atoms sticks) and the wt hTS (grey cartoon) in complex with FdUMP and raltitrexed (both in green and carbon atoms sticks) (**A**) or with the wt hTS (magenta cartoon) in complex with dUMP and raltitrexed (both in sticks, magenta carbon atoms) (**B**). The reduced pteridine moiety of 5-FHTF-di-Glu is moved with respect to the corresponding moiety of raltitrexed of both the ternary complexes, protruding in the substrate uracil site. In the structure of HT-hTS T251E, the catalytic cysteine is modified as S,S-(2-hydroxyethyl) thiocysteine (CME195, in grey carbon atoms sticks).

Former structural studies on hTS have reported a correlation between the enzyme active/inactive conformation and the ionic strength of the precipitant solutions used to crystallize the enzyme [123, 126]. Precipitant solutions having low concentrations (<0.2 M) of ammonium sulfate (low-salt condition) favor the switch of the enzyme to active conformation, whereas high concentrations (>1 M, high-salt condition) invariantly yielded the inactive conformation [58, 77, 126]. Our interface variants crystallized in the active conformation, despite the presence of ammonium sulfate in the precipitant solution. Indeed, both HT-hTS T251E and HT-hTS Q62R-T251E crystallized in sulfate-free conditions (see

Section 4.1.6, whereas the precipitant used for Q62R crystallization included high concentration of ammonium sulfate (2.0–2.2 M) [108]. This seems to suggest that the three interface variants are all stabilized in the active conformation, at least in part, by the population of the cofactor site by 5-FTHF or 5-FTHF-di-Glu. This observation also correlates with the results of the CD thermal denaturation analyses, showing an increment of the $\mathrm{T_m}$ for HT-hTS upon dUMP addition, whereas all interface variants retained the same $\mathrm{T_m}$ despite the presence of the substrate. Upon exposure to the substrate hTS is thought to switch to the active conformation, leading to increased thermal stability. On the other hand, all interface variants are already in the active conformation and unchanged thermal parameters are determined upon substrate addition. The structural data also contribute to explain the kinetic parameters reported for the interface variants (*Table 4.6*). All of them displayed increased $\mathrm{K_M}$ for mTHF, indicating a reduced affinity for the cofactor. This could be explained, at least in part, by the presence of folate-like molecules in the active site, that compete with the cofactor for its binding pocket. All variants also displayed lower turnover rates and reduced catalytic efficiencies. These reduced kinetic properties do not only account for the population of the cofactor site by exogenous folate-like compounds, but they are also reasonably due to the dimer destabilization effects induced by the interface mutations. Since both residues 62 and 251 are localized more than 15 Å away from the catalytic cavity, a direct effect of their mutation on ligand binding seems unlikely. Nevertheless, we can hypothesize long-range effects of the interface mutations on the catalytic activity or even on the active/inactive conformation equilibrium, peculiar of this enzyme. Former studies on hTS have shown that the introduction of mutations can affect the active/inactive equilibrium of the enzyme [61, 66]. Gibson and coworkers [66], reported that the hTS R163K variant is stabilized in the active conformation. Further investigations are needed to fully elucidate the effects induced by these interface point mutations on hTS and how they affect the active/inactive conformation equilibrium of the enzyme.

4.2.4 Structural characterization of the complexes HT-hTS Q62R-T251E:dUMP and HT-hTS Q62R-T251E:FdUMP

The crystal structure of HT-hTS double mutant Q62R-T251E was also solved in complex with the substrate dUMP and with its analogue FdUMP to resolutions of 2.8 Å and 2.6 Å, respectively (*Tables 4.3* and *4.5*). As for the dUMP-free structures, the catalytic Cys195 is exposed inside the catalytic cavity (active conformation). The analysis of electron density revealed the presence of dUMP and its analogue FdUMP bound to the substrate pocket in their respective complexes. Furthermore, the cofactor analogue 5-FTHF-di-Glu was also observed populating the cofactor binding site of both complexes. In the dUMP complex, the pyrimidine carbonyl oxygen in position 2 accepts two H-bonds from Asp218 backbone nitrogen and Gln214 amide nitrogen. The uracil N3 donates an H-bond to the Asn226 carbonyl group whereas the pyrimidine O4 is H-bonded to Asn226 amide and forms a water-mediated interaction with Glu87. The dUMP uracil C6 atom is placed ≈3.30 Å apart from the Cys195 Sγ, excluding the formation of a covalent bond. The hydroxyl of the dUMP ribose is H-bonded to His256 and Tyr258, whereas its phosphate moiety is salt bridged to Arg50, Arg215, Arg175', and Arg176' (*Figure 4.13*). The binding of FdUMP closely resemble that of the substrate, forming also conserved interactions inside the cavity. The FdUMP C6 atom is positioned slightly closer to the Cys195 Sγ, showing an average interatomic distance of ≈3.00 Å. FdUMP differs from the substrate by the 5-fluorine substituent on the pyrimidine ring, that takes additional halogen bond with Tyr135 hydroxyl (*Figure 4.14*). The binding mode of 5-FTHF-di-L-Glu is the same in both complexes and is also similar to the HT-hTS Q62R-T251E substrate-free structure (*Section 4.2.3.3*) . The major difference concerns the position of the reduced pteridine moiety of 5-FTHF di-L-Glu that in the structure of substrate-free HT-hTS Q62R-T251E is slightly shifted towards the substrates site, partially hindering it. In the ternary complexes with dUMP and FdUMP, the population of the substrate pocket induces a slight movement of the reduced pteridine moiety of

5-FTHF-di-L-Glu, placing the 5-formyl substituent ≈1 Å away with respect to the position occupied in the substrate free-structure. This allows to retain analogous interactions with the surrounding residues in all structural models. In the ternary complexes, additional interactions are entailed between the dUMP/FdUMP pyrimidine ring and the quinazoline ring of 5-FTHF-di-Glu, which are stacked to each other. The structures of our ternary complexes with the HT-hTS double mutant Q62R-T251E are also similar to other ternary complexes of hTS determined in presence of substrate and cofactor-like molecules, like raltitrexed [62, 126]. Only slight differences are observed in the cofactor-like molecules accounting for their different bicyclic systems.

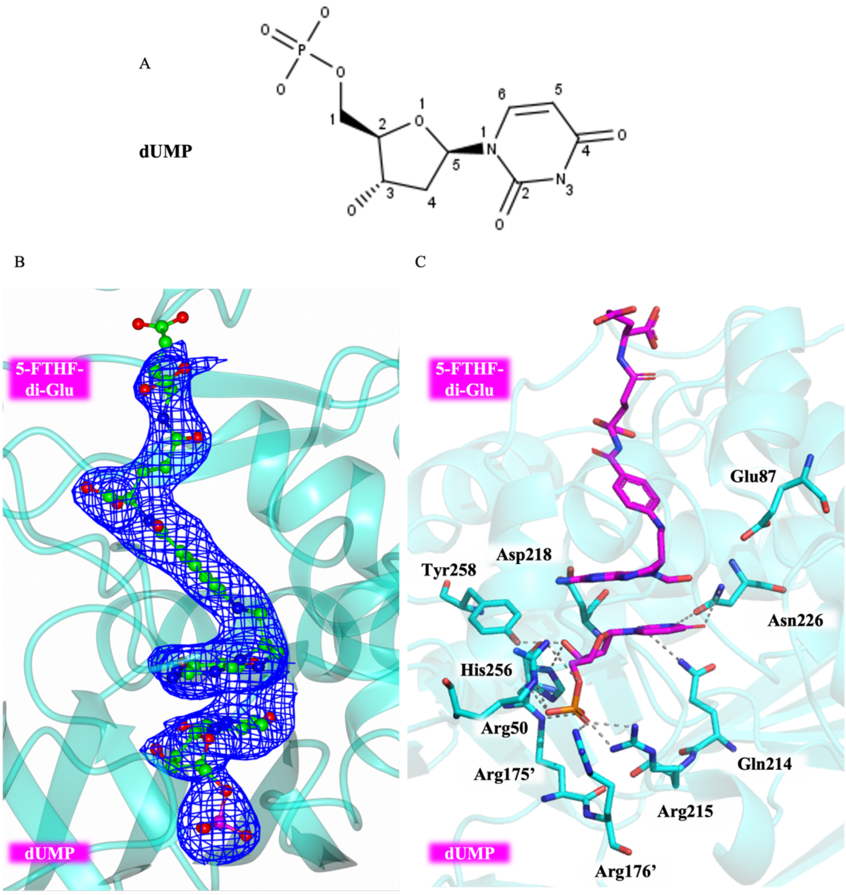


Figure 4.13: (**A**) Chemical structure of 2'-deoxyuridine-5'-monophosphate (dUMP). (**B**) Cartoon representation of active site view of HT-hTS Q62R-T251E:dUMP (in cyan cartoon). The substrate dUMP (in sticks, green carbons) is entrapped inside the catalytic cavity with the cofactor analogue 5-FTHF (in sticks, green carbons) both surrounded by the omit map contoured at the 2.5 σ level. dUMP and 5-FTHF are observed in all active sites of the four dimers found in the ASU (**C**) Cartoon representation of active site view of HT-hTS Q62R-T251E:FdUMP (in cartoon, interacting residues in sticks, cyan carbons). FdUMP takes several H-bonds (grey dashed lines) with the surrounding residues (Asp218, Gln214, Asn226, Glu87;His256,Tyr258; in sticks, cyan carbons), a water mediated interactions with Glu87, and salt-bridges (grey dashed lines) through its phosphate mojety with Arg50, Arg215, Arg175', and Arg176' (in sticks, cyan atoms). The cofactor analogue 5-FTHF-diGlu establishes the same tight network of H-bonds and van der Waals interactions visualized in *Figure 4.11* and here omitted for clarity.

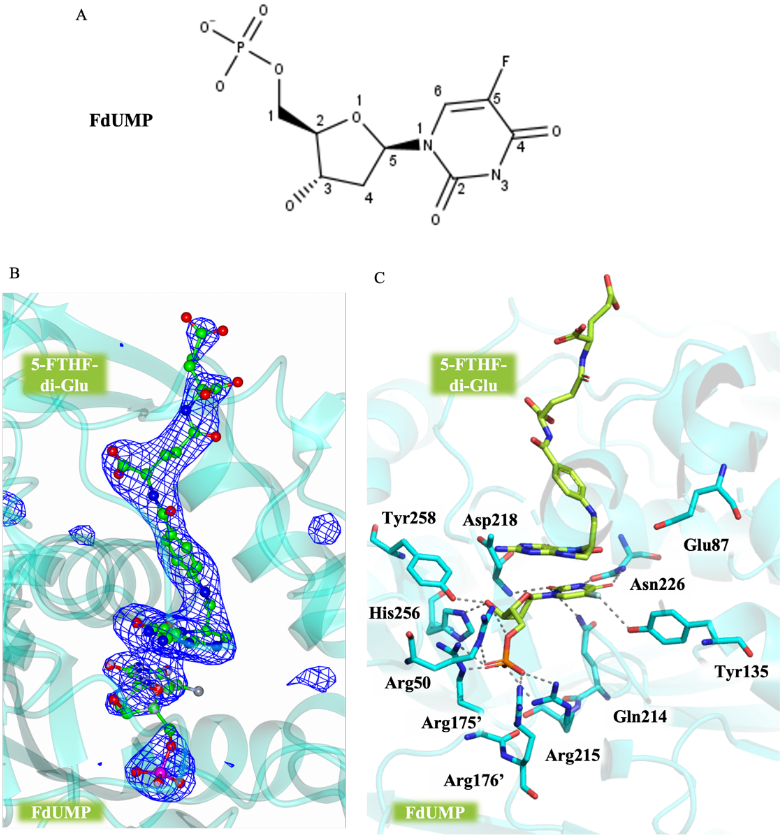


Figure 4.14: (**A**) Chemical structure of 5’-fluorodeoxyuridine-monophosphate (5-FdUMP). (**B**) Cartoon representation of active site view of HT-hTS Q62R-T251E:FdUMP (in cyan cartoon). The substrate analogue 5-FdUMP (in sticks, green carbons) is entrapped inside the catalytic cavity with the cofactor analogue 5-FTHF (in sticks, green carbons) both surrounded by the omit map contoured at the 2.5 σ level. FdUMP and 5-FTHF are observed in all active sites of the four dimers found in the ASU. (**C**) Cartoon representation of active site view of HT-hTS Q62R-T251E:FdUMP (in cyan cartoon, interacting residues in sticks, cyan carbons). FdUMP takes several H-bonds (grey dashed lines) with the surrounding residues (Asp218, Gln214, Asn226, Glu87;His256,Tyr258; in sticks, cyan atoms), a water mediated interactions with Glu87, and salt-bridges (grey dashed lines) through its phosphate mojety with Arg50, Arg215, Arg175’, and Arg176’ (in sticks, cyan atoms). The 5-fluorine substituent on the pyrimidine ring, takes halogen bond with the Tyr135 hydroxyl. The cofactor analogue 5-FTHF-diGlu establishes the same tight network of H-bonds and van der Waals interactions visualized in *Figure 4.11* and here omitted for clarity.

4.3 Evidence of destabilization of the human Thymidylate Synthase (hTS) dimeric structure induced by the interface mutation Q62R

Publication: Pozzi C, Lopresti L, Santucci M, Costi MP, Mangani S. Evidence of Destabilization of the Human Thymidylate Synthase (hTS) Dimeric Structure Induced by the Interface Mutation Q62R. Biomolecules. 2019 Apr 3;9(4):134.

Contribution: My contribution to this work was protein expression and purification, and C.D. thermal denaturation analysis.

Copyright: Article published under an open access Creative Common CC BY license, no special permission is required to reuse all or part of article published by MDPI.

biomolecules

Article

Evidence of Destabilization of the Human Thymidylate Synthase (hTS) Dimeric Structure Induced by the Interface Mutation Q62R

Cecilia Pozzi [1,*], Ludovica Lopresti [1], Matteo Santucci [2], Maria Paola Costi [2] and Stefano Mangani [1,*]

[1] Department of Biotechnology, Chemistry and Pharmacy—Department of Excellence 2018-2020, University of Siena, 53100 Siena, Italy; lopresti4@student.unisi.it

[2] Department of Life Sciences, University of Modena and Reggio Emilia, 41125 Modena, Italy; matteo.santucci86@gmail.com (M.S.); mariapaola.costi@unimore.it (M.P.C.)

* Correspondence: pozzi4@unisi.it (C.P.); stefano.mangani@unisi.it (S.M.); Tel.: +39-0577-232132 (C.P.); +39-0577-234255 (S.M.)

Received: 18 March 2019; Accepted: 1 April 2019; Published: 3 April 2019

Abstract: In human cells, thymidylate synthase (TS) provides the only source of 2′-deoxythymidyne-5′-monophosphate (dTMP), which is required for DNA biosynthesis. Because of its pivotal role, human TS (hTS) represents a validated target for anticancer chemotherapy. Nonetheless, the efficacy of drugs blocking the hTS active site has limitations due to the onset of resistance in cancer cells, requiring the identification of new strategies to effectively inhibit this enzyme. Human TS works as an obligate homodimer, making the inter-subunit interface an attractive targetable area. Here, we report the design and investigation of a new hTS variant, in which Gln62, located at the dimer interface, has been replaced by arginine in order to destabilize the enzyme quaternary assembly. The hTS Q62R variant has been characterized though kinetic assay, thermal denaturation analysis and X-ray crystallography. Our results provide evidence that hTS Q62R has a reduced melting temperature. The effective destabilization of the TS quaternary structure is also confirmed by structural analysis, showing that the introduced mutation induces a slight aperture of the hTS dimer. The generation of hTS variants having a more accessible interface area can facilitate the screening of interface-targeting molecules, providing key information for the rational design of innovative hTS interface inhibitors.

Keywords: human thymidylate synthase; interface variant; dimer destabilization; circular dichroism; thermal stability; X-ray crystallography; site-directed mutagenesis

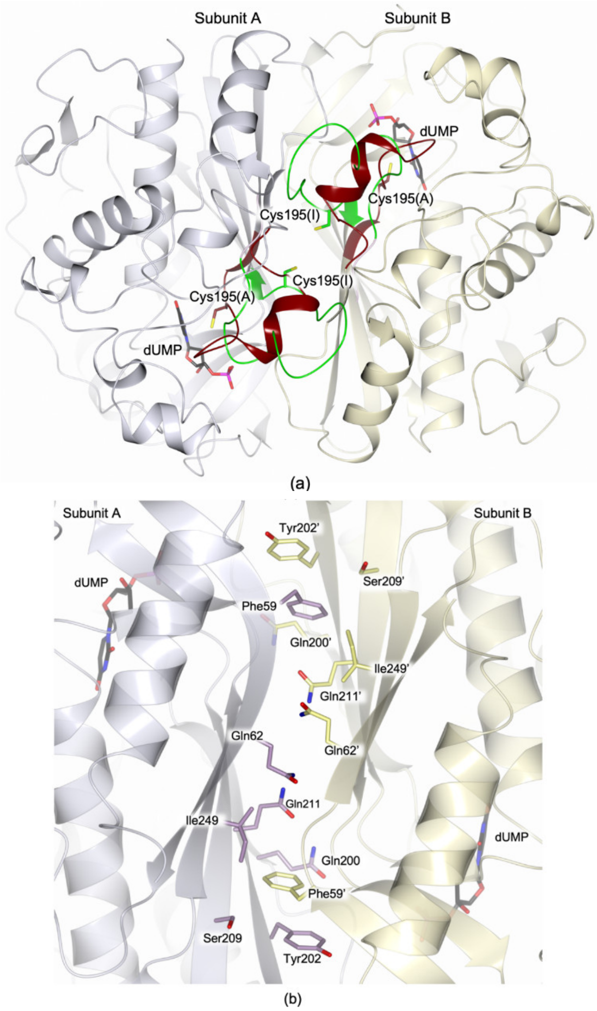
1. Introduction

Thymidylate synthase (TS, EC 2.1.1.45) plays a pivotal role in human cells, since it catalyzes the reductive methylation of 2′-deoxyuridine-5′-monophosphate (dUMP) to 2′-deoxythymidyne-5′-monophosphate (dTMP), using N^5,N^{10}-methylenetetrahydrofolate (mTHF) as cofactor. According to the reaction mechanism, the thiol of the catalytic Cys195 attacks the carbon atom in position 6 (C6) on the dUMP pyrimidine, forming a covalent adduct (Figure S1) [1]. The dUMP uracil carbon in position 5 (C5) is thus activated to accept the methyl moiety (C11) and the hydride donated by mTHF. In human cells, TS provides the only synthetic source of dTMP necessary for DNA biosynthesis, indeed its inhibition halts the replication processes and induces apoptosis in rapidly dividing cells, an effect known as “thymineless death” [2]. This classifies human TS (hTS) as an important target in anticancer chemotherapy. As a matter of fact, various TS inhibitors targeting the enzyme active site such as FdUMP (the active metabolite of 5-fluorouracil) and raltitrexed, are currently in use as anticancer agents [3,4]. Nonetheless, the use of these classical hTS inhibitors has restrictions due to the onset of resistance induced by TS overexpression [3–5]. Thus, new inhibition strategies have to be explored to effectively inhibit hTS without causing resistance in cancer cells.

Former structural studies have revealed that the hTS homodimers are able to switch between two alternate conformations, named active and inactive, primarily by changing the orientation of the loop including the catalytic Cys195 (catalytic loop, residues 181–197; Figure 1a) [6,7]. In the active conformation, Cys195 is exposed inside the catalytic cavity, whereas it is moved at the dimer interface upon the switch to the inactive form. The transition to the active conformation is fundamental to create the functional active site in which dUMP is accommodated, preceding the binding of cofactor (Figure S2) [7]. Inside cells, hTS acts also as a regulatory protein by binding RNAs, including its own mRNA (TSmRNA) [5,8]. It has been proposed that active site-targeting inhibitors of hTS stabilize the active conformation of the enzyme that reduces the affinity of TS for the TSmRNA [5]. This removes the translational arrest and triggers TS overexpression, leading to drug resistance [5]. The TSmRNA binding site on hTS is yet uncharacterized. Even so, there is evidence that the hTS dimer interface plays an important role in hTS-mRNA recognition, perhaps by controlling conformational transitions that alternatively expose and hide the TSmRNA recognition site [9–11]. The hTS dimer interface represents an attractive targetable area for the development of new inhibitors that could avoid the onset of drug resistance. Nonetheless, the generation of interface-targeting molecules is tricky and the identification of targetable spots on the interface area is a fundamental requisite. For this purpose, we formerly probed the hTS dimer interface through the generation of a set of alanine variants [12]. Among them, the mutant F59A showed a meaningful dimer destabilization effect, resulting in a K_d of 10 (±2) × 10^{-5} M, more than three orders of magnitude higher than that measured for the wild-type hTS, K_d of 5 (±1) × 10^{-8} M (both determined through fluorescence resonance energy transfer, FRET, based assays) [12]. This significant gain in the dissociation constant was interpreted as a shift of the monomer-dimer equilibrium towards the monomeric form [12]. This evidence further suggests that the area surrounding Phe59 (Figure 1b) is important to stabilize the enzyme quaternary structure, classifying it as a potentially druggable spot for the development of interface-targeting inhibitors.

On the basis of these results, we have introduced a bulky, charged residue at the dimer interface close to Phe59, in order to generate destabilized homodimers that can facilitate the access to this area, simplifying the identification of interface-binding molecules. The hTS variant Q62R, having the interface residue Gln62 replaced by arginine, was generated and characterized through kinetic assay, circular dichroism (CD) thermal denaturation analysis and X-ray crystallography. Residue 62 is proximal to Phe59 and it faces itself on the cognate subunit enhancing the electrostatic repulsion effects induced by the introduction of a charged residue in this position (Figure 1b).

The CD thermal denaturation studies, performed on the wild-type enzyme and the Q62R variant, showed meaningful variations in the melting temperature of the mutant, evidencing a destabilization effect induced by the introduction of Arg62 at the dimer interface. The structure of hTS Q62R was determined and compared to that of the wild-type enzyme, highlighting a slight aperture of the mutated enzyme homodimer. Furthermore, the structural analysis revealed that the variant adopts the active conformation, entrapping a cofactor analogue molecule within the catalytic cavity. The kinetic characterization of hTS Q62R is consistent with the crystallographic evidence. Our results show that the mutation Q62R effectively destabilizes the hTS homodimer, supporting the importance of this area for the enzyme quaternary assembly.



(a)

(b)

Figure 1. (**a**) Cartoon representation of the superimposition between the human thymidylate synthase (hTS) homodimer (subunits A and B are colored light lilac and yellow, respectively) in the active (Protein Data Bank, PDB, id 5X5D [13]) and inactive (PDB id 3N5G [14]) conformations. The two orientations of the catalytic loop (residues 181–197), defining the active (brown trace; PDB id 5X5D [13]) and inactive (green trace; PDB id 3N5G [14]) conformations, are displayed. The catalytic cysteine is shown in sticks in the active (A) and inactive (I) conformations (brown and green carbons, respectively). The position of the catalytic cavity is indicated by the presence of the substrate 2′-deoxyuridine 5′-monophosphate (dUMP, in sticks, black carbons; PDB id 5X5D [13]). (**b**) Interface view of the two Phe59 pockets, proved to be important for enzyme dimerization [12]. The position of the nearby Gln62, facing Gln62′ on the cognate subunit, is shown. Residues are displayed in sticks (carbon atoms are color-coded according to the parent subunits). In all figures, nitrogen atoms are colored blue, oxygen red, sulfur yellow, and phosphorous magenta.

2. Materials and Methods

2.1. Cloning and Site-Directed Mutagenesis

The hTS variant Q62R was generated by site-directed mutagenesis using partially overlapping primers (forward primer: GCATGCGGGCGAGATATTCATTAC; reverse primer: CTCGCCCGCATGCCGAAGACGC; purchased from Sigma-Aldrich, Milan, Italy) and the pQE80L-hTS plasmid (including the gene coding sequence for hTS cloned within the BamHI–HindIII restriction sites) as template. The 50 µL PCR reaction mixture included 50–100 ng template DNA, 1 µM primer pair, 200 µM dNTPs and Expand High Fidelity (Roche Biochemicals, Basel, Switzerland). The reaction was performed through an initial denaturation step (94 °C, 5 min), followed by 25 cycles of denaturation (94 °C, 1 min), annealing (52 °C, 1 min), and expansion (72 °C, 10 min), and by a final extension step (72 °C, 20 min). Afterwards, an aliquot (10 µL) of the reaction mixture was incubated overnight at 37 °C with 1 µL of DpnI (NEB, Ipswich, MA, USA). The resulting sample was heat-shock transformed in the *E. coli* TOP10 strain and positive transformants were selected on LB-Agar plates supplemented with 100 mg L^{-1} ampicillin. A single colony was cultured in LB medium added by 100 mg L^{-1} ampicillin (14 h, 37 °C, 220 rpm) and used for plasmid extraction (EZNA DNA extraction kit I, Omega Bioteck, Norcross, GA, USA). Site-directed mutagenesis was confirmed by sequencing the entire TS gene (the absence of unwanted mutations was also verified; sequencing service performed by Eurofins, Ebersberg, Germany).

2.2. Protein Expression and Purification

The wild-type hTS was expressed as His^6-tag protein (HT-hTS, the non-cleavable N-terminal His^6-tag was encoded by the pQE80L expression plasmid) in the *E. coli* strain BL21(DE3) as previously described [12], with minor modifications. Briefly, bacteria were cultured at 30 °C in the auto-induction medium ZYP-5052 [15] for 30 h. Cells, harvested by centrifugation (3000 g, 15 min, 8 °C), were resuspended in buffer A (50 mM HEPES, pH 7.5 and 30 mM NaCl), added by 20 mM imidazole, 0.2 mM phenylmethylsulfonyl fluoride (PMSF) and 0.5 mg mL^{-1} lysozyme, and then disrupted by sonication after 60 min incubation on ice. The cell-free extract, obtained by centrifugation (12,000 g, 60 min, 8 °C), was applied to a HisTrap HP 5 mL column (GE Healthcare, Milan, Italy) and eluted using a step-gradient protocol, by applying 250–500 mM imidazole concentration in the same buffer. Fractions containing the target protein were pooled and dialyzed in buffer A. The resulting sample was concentrated (Vivaspin 20 molecular weight cut-off 10 kDa, Sartorius, Göttingen, Germany) and further purified by size exclusion chromatography on a HiLoad 16/600 Superdex 75pg column (GE Healthcare, Milan, Italy). The elution profile was consistent with the enzyme dimer assembly (not shown). The high purity (>98%) of the resulting protein sample was confirmed by SDS-PAGE analysis (NuPAGE 4-12% Bis-Tris protein gels; Thermo Fisher Scientific, Waltham, MA, USA) and MALDI-TOF mass spectrometry (Toscana Life Science, Siena, Italy).

Attempting to produce the HT-hTS Q62R variant under the condition optimized for the wild-type enzyme resulted in an almost complete localization of the target protein in inclusion bodies. Therefore, a wide set of expression conditions was screened by testing different culturing media (Luria Broth, Super Broth, ZYP-5052), incubation temperatures (20 and 30 °C), inductor concentrations (isopropyl β-D-thiogalactopyranoside, IPTG, 0.1 and 0.5 mM), and incubation times (20 and 48 h). Our best condition turned out to be by culturing bacteria (plasmid transformants of *E. coli* BL21(DE3)) in the auto-induction medium ZYP-5052 [15] at 20 °C for 48 h. The HT-hTS Q62R variant was purified by nickel-affinity and size exclusion chromatography following the same procedure described for the wild-type enzyme.

2.3. Enzymatic Activity Assays

Enzyme activity assays were performed spectrophotometrically, according to a reported protocol [12]. Briefly, 1 mL reaction mixtures were prepared by adding aliquots of the enzyme

(0.14–1.50 μM) to the assay buffer (50 mM TES, pH 7.4, 25 mM $MgCl_2$, 6.5 mM HCHO, 1 mM EDTA, 75 mM β-mercaptoethanol) including variable concentrations of dUMP (3–180 μM) and mTHF (5–75 μM). Reactions, started by the addition of the substrate, were monitored by following the increase in absorbance at 340 nm during the oxidation reaction of mTHF to 7,8-dihydrofolate (DHF), for 3 min. K_M (Michaelis-Menten constant) values were determined for both mTHF and dUMP by varying their concentration in the assays, whereas k_{cat} was determined by changing the enzyme concentration.

2.4. Circular Dichroism (CD) Thermal Denaturation Analysis

The thermal stability of wild-type HT-hTS and the Q62R variant was evaluated using thermal unfolding experiments by monitoring the far-UV CD signal at 220 nm, on a Jasco (Pfungstadt, Germany) J-815 spectropolarimeter. The protein samples (20 μM enzyme solution in 10 mM HEPES pH 7.5 and 100 mM NaCl) were heated from 25 to 75 °C at a rate of 1 °C min^{-1}. Two sets of experiments were performed on the HT-hTS and the Q62R variant. The first was performed on the purified proteins, whereas the second on the samples exposed to 1 mM dUMP for 30 min on ice. Measurements were performed in triplicates. Data were analyzed using the software GraphPad Prism 7 in non-linear regression using the Boltzmann sigmoidal function for melting temperature (T_m) determination.

2.5. Protein Crystallization

Prior to crystallization experiments the purified protein was concentrated to 20 mg mL^{-1} and stored at −20 °C (until required). Crystallization trials were performed on the purified HT-hTS Q62R (20 mg mL^{-1} in 50 mM HEPES pH 7.5, 30 mM NaCl, with or without 20 mM β-mercaptoethanol) using the commercially available kits PEG/Ion, Index and Grid screen Ammonium Sulfate from Hampton Research (Aliso Viejo, CA, USA), and JBScreen Basic (JBSB) 1–4 and Classic (JBSC) 6 from Jena Bioscience (Jena, Germany). Protein crystal growth was observed in 10–14 days using the JBSC6 solution C4 (2.0 M ammonium sulfate, 100 mM TRIS pH 8.5) as precipitant.

The optimization of the crystallization condition was performed using the hanging drop vapor-diffusion method [16] at 20 °C, by varying both the ammonium sulfate concentration and the buffer. Crystals, suitable for diffraction experiments, were obtained from drops prepared by mixing equal volumes of protein (above protein solution, with or without 10 mM dUMP) and precipitant (2.0–2.2 M ammonium sulfate, 100 mM bicine pH 9.0) solutions, equilibrated over 800 μL reservoir. Crystal growth was observed within two weeks only in drops prepared by including 20 mM β-mercaptoethanol in the sample solution. Prior to X-ray diffraction experiments, crystals were transferred to the cryoprotectant solution (20% vol/vol glycerol, 2.4 M ammonium sulfate, 100 mM bicine pH 9.0) and flash frozen in liquid nitrogen.

2.6. Data Collection, Structure Solution and Refinement

X-ray crystallographic data were collected using synchrotron radiation at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) beamline ID30B, equipped with a Dectris (Baden-Daettwil, Switzerland) Pilatus3 6M detector. Reflections were indexed and integrated using the program XDS [17] and scaled with SCALA [18] from the CCP4 suite [19]. Data collection and reduction statistics are displayed in Table S1. Crystals of HT-hTS Q62R belonged to the primitive orthorhombic space group $P2_122_1$, including eight enzyme subunits (four enzyme dimers) in the cell asymmetric unit (ASU). The structure was solved by molecular replacement using the software Molrep [20] from the CCP4 suite. One monomer of hTS in the active (PDB id 1HVY [7]) and inactive (PDB id 3N5G [14]) conformations were attempted as searching models (excluding water molecules and non-protein atoms), providing clear evidence that the enzyme crystallized in the active conformation (active conformation: score of 0.775 and wRfac of 0.423; inactive conformation: score of 0.599 and wRfac of 0.548). The structure was refined with Refmac5 [21] from the CCP4 suite using the TLS parametrization [22] in the last cycles of refinement. The optimal partitioning of the polypeptide chains was calculated though the *TLS Motion Determination* web server [23], resulting in twenty continuous segments. The molecular graphic software Coot [24,25] was used for manual rebuilding and modelling of missing atoms. Water molecules were added through the ARP/wARP suite [26] and checked with Coot. Upon completion of the protein model, inspection of the Fourier difference map clearly evidenced the presence of a ligand bound in the active site of all enzyme subunits. The shape of the map indicated that the ligand was a derivative of tetrahydrofolate (THF) modified at position 5 of the pteridine ring. The two THF derivatives 5-formyl-6-tetrahydrofolate (5-FHTF) and 5-hydroxymethyl-6-tetrahydrofolate (5-HMTHF) were alternatively modelled and refined in this site (in all active sites of the four enzyme dimers). Furthermore, sulfate anions and glycerol molecules from crystallization/cryoprotectant solutions were found within both enzyme dimers (seventeen sulfate anions and two glycerol molecules were collectively included in the model). The occupancies of the exogenous ligands were singularly adjusted to values, resulting in atomic displacement parameters close to those of neighboring protein atoms in fully occupied sites. The stereochemical quality of the final model was checked using Coot and Procheck [27]. Refinement statistics are reported in Table S2. Figures were generated through the molecular-graphic software CCP4mg [28].

2.7. Protein Data Bank (PDB) Deposition

Atomic coordinates and structure factors for HT-hTS Q62R were deposited in the Protein Data Bank under the accession code 6R2E.

3. Results

3.1. Variant Production

The hTS Q62R variant was generated through site-directed mutagenesis, using partially overlapping primers and the gene coding sequence for the wild-type enzyme as template for the PCR reaction. In the resulting amplified DNA, the gene coding sequence for the hTS Q62R variant was inserted in the pQE80L vector (plasmid pQE80L—hTS-Q62R), which also included the coding sequence for a non-cleavable N-terminal His^6-tag. The variant was expressed as His^6-tag protein (HT-hTS Q62R) in the bacterial strain *E. coli* BL21(DE3). Attempting to express the variant under the same experimental conditions adopted for the wild-type enzyme resulted in the almost complete localization of the target protein in inclusion bodies. To improve the solubility of HT-hTS Q62R, a wide set of expression conditions was screened relying on different incubation temperatures, culture media, inductor concentrations, and incubation times. As expected, the formation of inclusion bodies was generally decreased by reducing the incubation temperature. Indeed, our best conditions turned out to be caused by culturing bacterial cells at 20 °C in the ZYP-5052 auto-induction medium for 48 h.

The purification procedure took advantage from the introduction of the N-terminal His^6-tag, indeed almost pure (>95%) protein samples for HT-hTS and the Q62R variant were obtained after the first purification step relying on nickel-affinity chromatography. The purification was completed through size exclusion chromatography, resulting in highly pure protein samples (>98%). The final production yield for HT-hTS Q62R was estimated to ~80 mg L^{-1}. In contrast, yields of ~200 mg L^{-1} were reported for the wild-type enzyme under standard expression conditions [12,29], and confirmed by us through our modified expression protocol (yield of ~250 mg L^{-1} using the ZYP-5052 auto-induction medium).

3.2. Kinetic Characterization

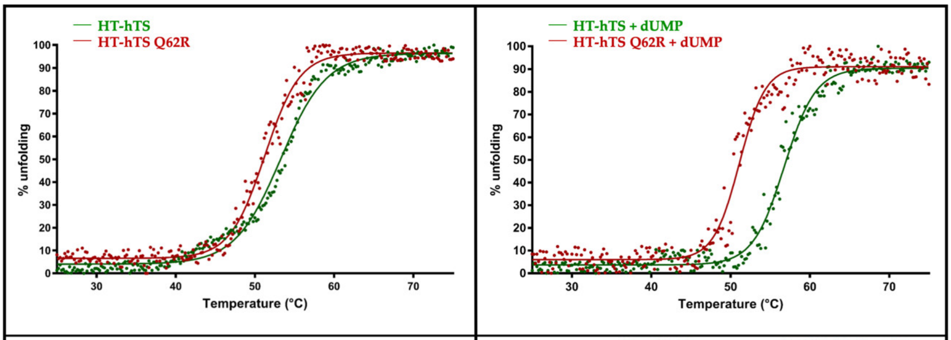
The enzymatic activity assays performed on the HT-hTS and the Q62R variant showed that the point-mutation introduced at the dimer interface perturbs the kinetic properties of the enzyme (Table 1). For the variant Q62R, the K_M value for the substrate dUMP is almost unaltered with respect to HT-hTS, whereas the K_M value for the cofactor mTHF is ~2.5 times higher. More pronounced is the effect on the turnover rate (k_{cat}) that results decreased by more than 6 times in the Q62R variant. This reflects also on the catalytic efficiency, indeed the k_{cat}/K_M values determined for the variant are 5–17 times lower than those determined for the wild-type enzyme. These results are explained by the binding of a cofactor-analogue molecule to HT-hTS Q62R and by the destabilization effects induced on the dimer quaternary assembly (*vide infra*, Sections 3.3 and 3.4).

Table 1. Kinetic characterization of His^6-tag human thymidylate synthase (HT-hTS) and of its interface variant Q62R.

	K_M (dUMP) (μM)	**K_M (mTHF) (μM)**	**k_{cat} (s^{-1})**	**k_{cat}/K_M (dUMP) (μM^{-1} s^{-1})**	**k_{cat}/K_M (mTHF) (μM^{-1} s^{-1})**
HT-hTS	10 ± 1	6 ± 1	1.00 ± 0.01	1.00×10^{-7}	1.67×10^{-7}
HT-hTS Q62R	8 ± 1	16 ± 1	0.16 ± 0.06	0.20×10^{-7}	0.10×10^{-7}

3.3. Circular Dichroism (CD) Thermal Denaturation Analysis

In hTS, residue 62, either a glutamine, in the wild-type enzyme, or an arginine, in the Q62R variant, is exposed at the dimer interface on which it faces the same residue (either Gln62′ or Arg62′) on the cognate subunit. The effect of the introduction of a bulky charged residue at the dimer interface was evaluated through thermal denaturation analysis by monitoring the CD signal at 220 nm (Figure 2). The values of the melting temperature (T_m) determined for the purified HT-hTS and the Q62R variant resulted of 53.27 (±0.09) and 51.26 (±0.08) °C, respectively (Figure 2a). The decrease by ~2.0 °C in the T_m of the variant indicates that the mutation introduced at the enzyme dimer interface induces a destabilization of the protein. The drop in the T_m is even more evident after the addition of the substrate (to a 50-fold molar excess with respect to the protein concentration, Figure 2b). At variance with the Q62R variant, for which the same T_m (51.11 ± 0.10 °C) was determined after the substrate addition, the wild-type enzyme undergoes to a stabilization effect in presence of dUMP that increases its T_m to 56.77 (±0.08) °C (gain by ~3.5 °C). The comparison between the Q62R variant and the wild-type enzyme exposed to the substrate, evidences a reduction by ~5.7 °C in the T_m of the interface mutant (Figure 2b). Furthermore, the thermal denaturation profile of HT-hTS suggests a biphasic transition that is no longer observed after the dUMP addition and in the curves of the variant.



	HT-hTS	**HT-hTS Q62R**		**HT-hTS + dUMP (1 mM)**	**HT-hTS Q62R + dUMP (1 mM)**
T_m (°C)	53.27 ± 0.09	51.26 ± 0.08	T_m (°C)	56.77 ± 0.08	51.11 ± 0.10

(a) (b)

Figure 2. Thermal unfolding transition curves of HT-hTS and HT-hTS Q62R followed by circular dichroism. Two sets of curves were determined, one on the purified proteins (**a**) and the second on the samples incubated for 30 min with 1 mM dUMP (**b**). Melting temperature (T_m) values determined for the transitions are tabulated in the panels. Measures were performed in triplicate. The thermal denaturation profile of the wild-type HT-hTS suggests a biphasic transition, no longer observed following dUMP addition and in the curves of the variant.

3.4. Structural Characterization of the HT-hTS Variant Q62R

The structure of HT-hTS Q62R was solved to 2.55 Å resolution (Table S1), showing that the mutated protein retains the constitutive dimeric quaternary structure of the wild-type enzyme (Figures 3 and 4a). Four enzyme homodimers (A–B, C–D, E–F, G–H, in our model) were found in the cell asymmetric unit (ASU, Figure 3), all fully traced apart for the first twenty-five N-terminal residues (further the twelve residues belonging to the non-removable His^6-tag). From the initial phases of structure solution and refinement, it was evident that HT-hTS Q62R was in the active conformation. The four dimers are nearly identical, as testified by the root mean square deviation (rmsd) upon Cα matching that ranged from 0.15 Å to 0.77 Å among all enzyme subunits. The most evident differences are localized in subunits B and D, in which the N-terminal segments point in a distinct direction with respect to other subunits, as evidenced by the structural comparison displayed in Figure 3b. The maximal displacement, resulting in ~9.5 Å, is observed on the N-terminal Pro26 (measured between the Cα atoms).

3.4.1. The HT-hTS Q62R Active Site

The mutation introduced at the enzyme dimer interface did not affect the architecture of the active site, indeed it results fully consistent with formerly reported models for the active conformation of the enzyme [7,13]. Nonetheless, the analysis of the electron density in the active site area evidenced that the catalytic Cys195 was modified as *S*,*S*-(2-hydroxyethyl)thiocysteine (CME195, Figure 4b) by the reaction with β-mercaptoethanol, added to the protein sample prior to the crystallization experiments. The presence of this reducing agent was found to be critical for protein crystallization, indeed attempting to crystallize HT-hTS Q62R without β-mercaptoethanol (or with a different reducing agent) invariantly failed in crystal growth (as formerly observed also for the wild-type enzyme, unpublished results). In all subunits, a sulfate anion was observed nearby the catalytic cysteine, anchored to the guanidinium moieties of the four arginine residues Arg50, Arg215, Arg175′, and Arg176′ (the last two from the partner subunit, Figure 4b). These arginines are also responsible for the recognition of the dUMP phosphate moiety in the substrate binding site [7,13], mimicked here by the sulfate anion. Furthermore, in all active sites, the presence of a folate-like molecule was systematically observed (Figure 4a,b). The shape of the electron density evidenced a bent conformation of the pyrazine ring of the folate pteridine moiety peculiar to the cofactor reduced form, strongly suggesting that the ligand was a tetrahydrofolate (THF) derivative (Figure 4b). The substituent on the pteridine C5 was a bi-atomic species, consistent with either a hydroxymethyl moiety or a formyl group (5-ethyl derivatives of the cofactor are not known). The ligand was refined (in all active sites of the four dimers) either as 5-hydroxymethyl-6-tetrahydrofolate (5-HMTHF) or 5-formyl-6-tetrahydrofolate(5-FTHF), without meaningful changes in the refinement quality indicators (as expected) and in the resulting Fourier maps. Moreover, the resolution of the structure (2.55 Å) did not allow us to distinguish between single and double C-O bonds, preventing further speculations on which THF-derivative is observed in this site. Nonetheless, we opted for 5-FTHF because this molecule is naturally formed inside cells [30] and its complex with the bacterial *Enterococcus faecalis* TS (*Ef*TS, PDB id 3UWL) was formerly reported [31]. The observation of 5-FTHF is further suggested by the orientation, in all enzyme subunits, of the 5-formyl oxygen that is not engaged in intramolecular interaction with the adjacent carbonyl of the ketone group on the reduced pteridine. At variance with 5-FTHF, the presence of 5-HMTHF inside cells has never been detected [30]. We have attempted to refine the putative 5-HMTHF, just for the sake of completeness.

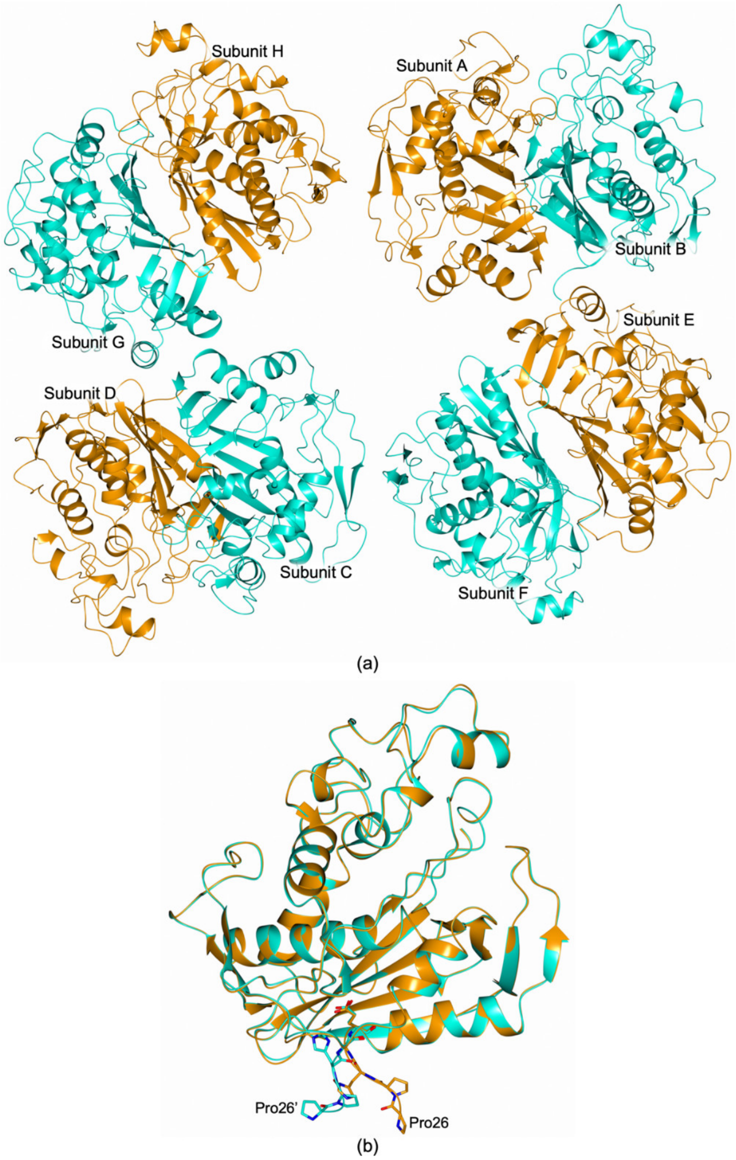
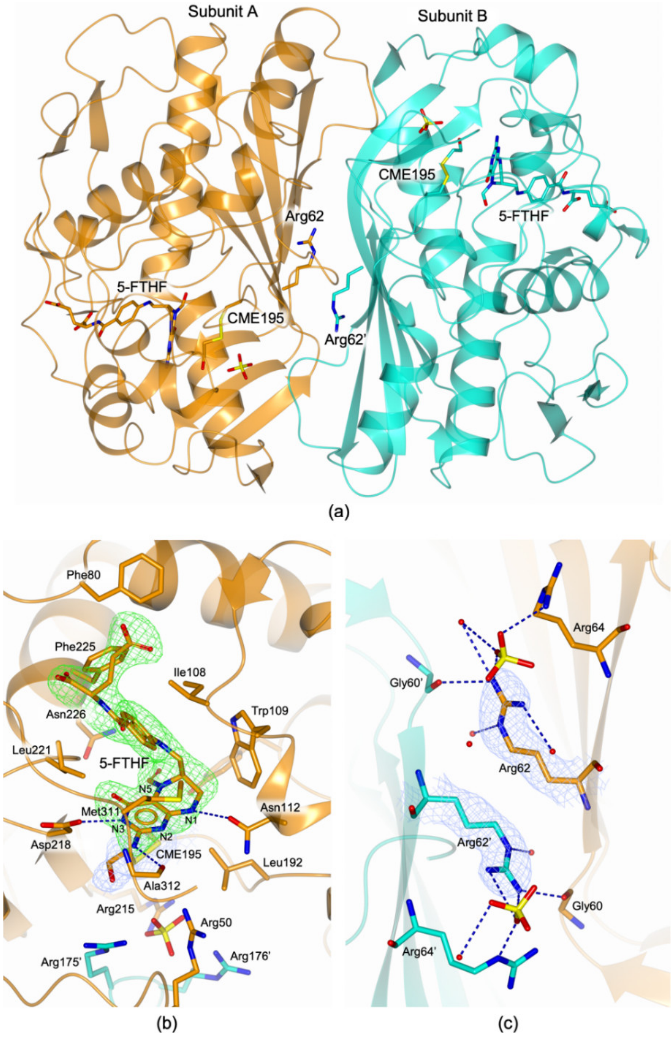


Figure 3. (**a**) Cartoon representation of the four HT-hTS Q62R homodimers (A–B, C–D, E–F, and G–H in our model) found in the cell ASU. (**b**) The structural comparison between subunit A and B (orange and cyan cartoon, respectively) shows that their N-terminal segments point in two distinct directions (residues 26–30 are shown in sticks, carbon atoms are color-coded according to the parent subunit). The maximal displacement, resulting of ~9.5 Å, is observed between Pro26 of two partner subunits (measured between their Cα atoms).



(a)

(b) (c)

Figure 4. (**a**) Cartoon representation of an enzyme homodimer (subunit A and B are colored orange and cyan, respectively). Within each enzyme homodimer, both subunits assume the active conformation, showing the catalytic Cys195 exposed inside the catalytic cavity. The catalytic Cys195 is modified as *S*,*S*-(2-hydroxyethyl)thiocysteine (CME195, in sticks) in all subunits. The mutated residue Arg62 is displayed in sticks. The cofactor analogue 5-formyl-6-tetrahydrofolate (5-FTHF, in sticks) is entrapped in the active site of all enzyme subunits. (**b**) Active site view of HT-hTS Q62R (in cartoon, interacting residues in sticks; subunit A and B are colored orange and cyan, respectively). The catalytic Cys195 is modified as *S*,*S*-(2-hydroxyethyl)thiocysteine (CME195, in sticks), as visible in the $2F_o$-F_c electron density map contoured at the 1.5 σ level. The cofactor analogue 5-FTHF (in sticks, orange carbons) is entrapped inside the catalytic cavity by a tight network of H-bonds (blue dashed lines) and van der Waals interactions (water molecules and water mediated interactions have been omitted for clarity). The ligand is surrounded by the omit map contoured at the 3 σ level. 5-FTHF is observed in all active sites of the four dimers found in the ASU. (**c**) Interface view of HT-hTS Q62R (in cartoon, interacting residues in sticks; subunit A and B are colored orange and cyan, respectively). The mutated Arg62 is surrounded by the $2F_o$-F_c electron density map contoured at the 1.5 σ level. Electrostatic interactions are displayed as blue dashed lines. Water molecules are shown as red spheres. Sulfate anions are displayed in sticks in all panels.

Within the HT-hTS Q62R active site, the cofactor analogue entails a tight network of H-bonds and van der Waals interactions (Figure 4b, only direct H-bonds are shown, water mediated interactions are omitted for clarity). The ketone moiety on the reduced pteridine ring forms water mediated interactions with Asp218, Asn226, and Gln214. Furthermore, the carboxylate moiety of Asp218 is positioned ~2.9 Å away from the pteridine nitrogen N3, strongly suggesting that it is protonated, while donating a H-bond to the protein residue (Figure 4b). The amine moiety on the reduced pteridine forms either direct or water mediated interactions with Ala312 and Asp218 (only the direct H-bond with Ala312 is shown in Figure 4b). The pteridine nitrogen N1 donates a H-bond to Asn112 (Figure 4b). Nearby, the nitrogen N2 forms water mediated interactions with the same Asn112 and with Arg50 (not shown in Figure 4b). Furthermore, 5-FTHF forms close van der Waals contacts with Ile108, Trp109, Leu192, Leu221, Phe225, and Met311.

3.4.2. The Arg62 Pocket

Residue 62 is localized at the periphery of the dimer interface in which it faces itself on the cognate subunit (Figure 1b). The shape of the electron density surrounding residue 62 clearly indicated the presence of an arginine in this site (Figure 4c). At the dimer interface, the side chains of the two facing Arg62 and Arg62′ are oriented in two opposite directions (Figure 4c). This arrangement is adopted to reduce the electrostatic repulsions induced by the presence of two facing charged residues. In this configuration, both arginines interact with the backbone carbonyl of Gly60 on the partner subunit and with various water molecules (Figure 4c). Furthermore, on the two dimer halves, Arg62 is directed towards Arg64 (belonging to the same subunit), creating a positively charged pocket on the enzyme surface in which a sulfate anion (deriving from the crystallization solution) is bound (Figure 4c). The distances separating the Cα of the two facing Arg62 ranged from 8.05 Å to 8.20 (±0.46) Å in the four enzyme dimers found in the ASU.

4. Discussion

4.1. 5-FTHF Binding in the HT-hTS Q62R Active Site

Former structural studies performed on hTS invariantly reported the yield of the enzyme in the inactive conformation in crystals grown under high-salt crystallization conditions (using precipitant solution including 1.0–1.4 M ammonium sulfate) [7,14,32]. In contrast, HT-hTS Q62R crystallized in the active conformation using a precipitant solution including a higher concentration of ammonium sulfate (2.0–2.2 M). The active conformation adopted in the crystal by HT-hTS Q62R can be ascribed to the population of the cofactor site by 5-FTHF (the formation of the analogous complex in the wild-type enzyme has never been reported). However, the point-mutation introduced at the dimer interface is unlikely to affect the ligand binding since Arg62 is localized more than 15 Å away from the catalytic cavity. Nonetheless, long-range effects have been observed in other enzymes such as *E. coli* class Ia ribonucleotide reductase [33]. On the other hand, the presence of 5-FTHF in the Q62R variant, contributes to explain the observed increase in the K_M of mTHF (Table 1), suggesting that the reduced affinity of the cofactor for its site is due to the binding of the ligand. The increased K_M measured for mTHF in the Q62R variant respect to HT-hTS indicates competition between cofactor and 5-FTHF, consistently with the crystallographic observation of the enzyme being in the active conformation with both subunits occupied by the cofactor analogue. The binding of 5-FTHF contributes also to explaining the reduced catalytic efficiency of the variant, together with the dimer destabilization effect induced by the mutation.

The wt-hTS—5-FTHF adduct in not available for comparison, but the complex of 5-FTHF with the bacterial *Ef*TS (PDB id 3UWL) has been previously characterized [31]. The comparison with HT-hTS Q62R shows that the ligand adopts the same pose in both enzymes (Figure 5a). Indeed, the active sites of *Ef*TS and hTS are widely conserved, the main difference being on the hTS residue Asn112 that is replaced by Trp84 in the bacterial enzyme (Figure 5a). Even so, both residues are involved in interactions with 5-FTHF, relying on the formation of either a H-bond in HT-hTS Q62R or van der Waals interactions in *Ef*TS (Figure 5a). In both structures, the catalytic cysteine (either Cys195 in HT-hTS Q62R or Cys197 in *Ef*TS) are modified by the reaction with β-mercaptoethanol added during the protein purification/crystallization procedures [31]. It is worth noting that in the structure of HT-hTS Q62R (and *Ef*TS), 5-FTHF populates the cofactor pocket independently by the binding of the substrate. To date all the structures reported for hTS in complex with cofactor analogue inhibitors have been determined in presence of the substrate that populates its site (ternary complexes) [7,13,34–36]. Our attempts to characterize HT-hTS Q62R in complex with the substrate dUMP have been unsuccessful. This is reasonably explained by the >200 times higher concentration of sulfate anions present in the precipitant solution that compete with the substrate for the population of the phosphate recognition pocket. Indeed, sulfate anions are observed in this site in all enzyme subunits (Figure 4b). The comparison with the structure of the ternary complex hTS-dUMP-raltitrexed (PDB id 5X5Q [13]) evidences a somewhat different arrangement of 5-FTHF and raltitrexed within the cofactor site (Figure 5b). In the structure of the Q62R variant, the reduced pteridine moiety of 5-FTHF is shifted towards the substrates site, hindering the uracil binding pocket.

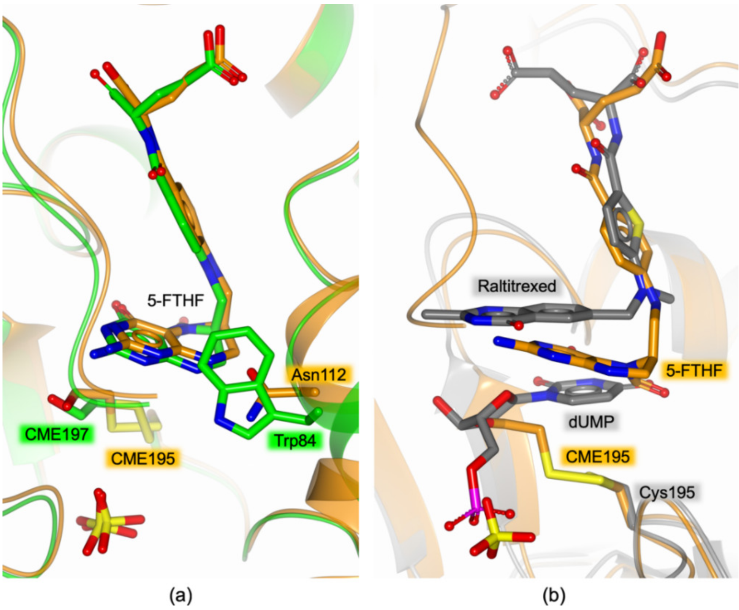


Figure 5. (**a**) Active site view of the superimposition between the structures of HT-hTS Q62R (orange cartoon and carbon atoms) and the bacterial *Enterococcus faecalis* TS (*Ef*TS, green cartoon and carbon atoms; PDB id 3UWL [31]). In both structures, the active site of the enzyme is populated by the cofactor analogue 5-formyl-6-tetrahydrofolate (5-FTHF, in sticks), showing a conserved binding mode. In both complexes, the catalytic cysteine is modified as *S*,*S*-(2-hydroxyethyl)thiocysteine (CME195 and CME197, in HT-hTS Q62R and *Ef*TS, respectively; in sticks). Sulfate anions are shown in sticks. (**b**) Active site view of the superimposition between the structures of HT-hTS Q62R (orange cartoon and carbon atoms, sulfate anion in sticks) in complex with 5-FTHF (in sticks) and the wild-type hTS (grey cartoon and carbons) in complex with 2′-deoxyuridine-5′-monophosphate (dUMP) and raltitrexed (both in sticks). The reduced pteridine moiety of 5-FHTF is moved with respect to the corresponding moiety of raltitrexed, protruding in the substrate uracil site. In the structure of HT-hTS Q62R, the catalytic cysteine is modified as *S*,*S*-(2-hydroxyethyl)thiocysteine (CME195, in sticks).

4.2. Effect of the Interface Point Mutation Q62R on the hTS Dimer Stability

Human thymidylate synthase works as an obligate dimer, showing a stable quaternary assembly due to an extended inter-subunit interface. Indeed, the analysis of the dimer interface through the PISA webserver (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html [37]) results in an average interface area of 2120.8 Å^2 and a $\Delta^i G$ of −20.4 kcal mol^{-1} ($\Delta^i G$, indicates the solvation free energy gain upon formation of the interface, calculated on the PDB id 5X5Q [13]). The same analysis performed on the structure of the Q62R variant shows a slight reduction of the interface area, resulting in 2032.0 Å^2, and a $\Delta^i G$ of −18.7 kcal mol^{-1}, suggesting that the introduced mutation induces a destabilization of the enzyme quaternary assembly. The negative effect on the enzyme stability is supported by the results obtained from thermal denaturation analysis. Indeed, for the Q62R variant, a drop in the T_m of ~2.0 °C is observed with respect to the wild-type, that increases to ~5.7 °C following the addition of the substrate dUMP (Figure 2). At variance with the Q62R variant, the wild-type enzyme is stabilized by the substrate, showing a gain in the T_m by ~3.5 °C (Figure 2). Former studies on hTS demonstrated that, in presence of the substrate, the active/inactive equilibrium (normally occurring in solution) shifts towards the active conformation in which dUMP is observed to bind [7,13]. Thus, the stabilization effect induced by dUMP on the wild-type enzyme is ascribed to the switch of hTS in the active conformation. Furthermore, the thermal denaturation profile observed for the wild-type HT-hTS, suggests that it occurs as a biphasic transition (Figure 2a), due to the active/inactive equilibrium, whereas a monophasic transition, due to the switch of the enzyme in the active conformation, is visible after the addition the dUMP (Figure 2b). Our results confirm those formerly obtained by Chen et al. through differential scanning fluorimetry analysis [13]. On the other hand, an analogous effect is not observed for the Q62R variant for which the same curves have been determined regardless the presence of the substrate. The behavior of the variant is explained by the structural characterization of the "as prepared" HT-hTS Q62R showing that the enzyme adopts the active conformation, stabilized by the interaction with 5-FTHF, that populates the active site. Even though the dUMP binding in the HT-hTS Q62R active site is allowed (indeed the variant is catalytically active, Table 1), the enzyme is already shifted in the active conformation (by the interaction with 5-FTHF), minimizing the stabilization effect induced by the dUMP addition.

The CD thermal denaturation experiments do not provide direct evidence about the structural reason of HT-hTS Q62R destabilization. However, the comparison of the mutant crystal structure with that of the wild-type enzyme in the active conformation suggests a likely explanation for the observed HT-hTS Q62R destabilization occurring in solution. The superimposition, displayed in Figure 6, provides clear evidence that the point-mutation Q62R induces a slight aperture of the enzyme dimer which is particularly visible in the area surrounding Arg62 and Arg62′ on the two enzyme halves. In the Q62R, a distance of ~8.1 Å is measured between the Cα of the two facing Arg62, whereas Gln62 and Gln62′ in the wild-type enzyme are placed ~6.6 Å apart. The local shift, by ~1.5 Å, is spread over a large interface area (Figure 6), inducing a weakening of the dimer quaternary assembly of the Q62R variant. This is also consistent with the slight reduction of the interface area observed by PISA analysis. The drop (5.7 °C) of the T_m of the variant is reasonably explained by the destabilization of the enzyme quaternary structure indicated by the increased inter-subunit distance induced by the mutation.

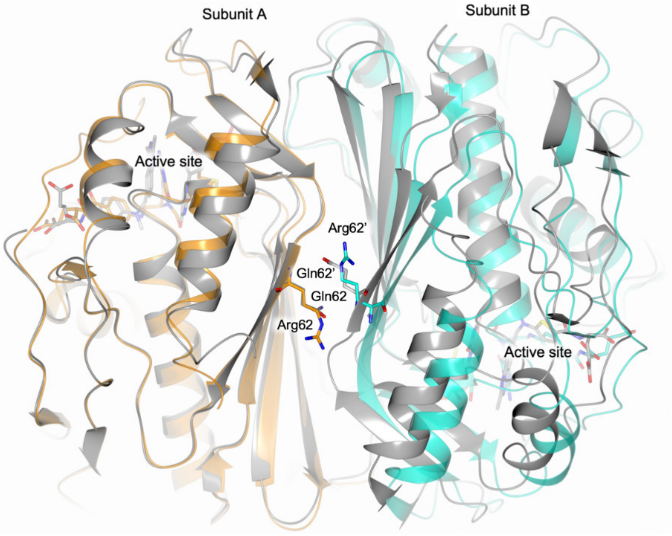


Figure 6. Structural comparison between the homodimer of HT-hTS Q62R (subunit A and B are colored orange and cyan, respectively) and the wild-type enzyme (in grey; PDB id 5X5Q [13]). The superimposition (performed by matching subunit A of both models) indicates a shift of the cognate subunit in the structure of the Q62R variant.

To confirm this explanation, we will attempt to perform FRET and NMR experiments in solution on the Q62R variant to ascertain the influence of this mutation on the monomer/dimer equilibrium (i.e., determination of the K_d and of the equilibrium shift towards the monomeric form).

5. Conclusions

Protein-protein interfaces (PPIs) are fundamental for the acquisition of the quaternary structure and for the interaction with partner proteins, but the development of interface targeting inhibitor is difficult. Even though the extension of the interface area is a challenge, the introduction of single-point mutation can induce meaningful perturbations that are exploitable for the development of interface-perturbing drugs. Obligate homodimer enzymes, as hTS, represents a special class of PPIs where the disruption of interface interactions abolishes their catalytic activity. Thus, the hTS dimer interface represents an attractive targetable area for the development of innovative hTS inhibitors. Here, we have investigated the hTS variant Q62R in which the interface residue Gln62 has been replaced by a bulkier charged arginine. The effect of the point mutation has been evaluated through kinetic analysis, CD thermal denaturation studies and X-ray crystallography, providing evidence that the mutation Q62R induces a destabilization effect on the enzyme dimeric structure. Furthermore, our results support the importance of this interface area for the dimer quaternary assembly, in agreement with former studies highlighting the key contribution provided by the nearby residue Phe59 [12]. Indeed, a slight aperture of the hTS dimer is observed subsequently to the integration of Arg62 at the periphery of the inter-subunit interface. The generation of homodimers having a slightly opened dimeric structure, such as the Q62R mutant, can facilitate the access of small molecules to the interface area, simplifying the screening of interface-targeting molecules. Validation procedures on homodimers of the wild-type enzyme are required to verify the effectiveness of the interface-directed molecules and to avoid false-positive binders (e.g., Arg62-interacting molecules). Thus, the hTS Q62R variant, may represent a functional tool exploitable to identify innovative interface-targeting inhibitors.

Supplementary Materials: The following are available online at http://www.mdpi.com/2218-273X/9/4/134/s1, Figure S1: Scheme for the TS catalyzed reaction and chemical structures of substrate and cofactor. Figure S2: Active site view of hTS. Table S1: Data collection and processing, Table S2: Structure solution and refinement.

Author Contributions: Variant generation, C.P.; Protein expression and purification, C.P. and L.L.; enzyme activity assays, M.S.; C.D. thermal denaturation analysis, L.L.; protein crystallization, crystallographic data collection, and structure solution and refinement, C.P.; writing and original draft preparation, C.P., S.M.; writing, review and editing, M.P.C., C.P., S.M.; Funding acquisition, M.P.C.

Funding: This work received final support by the Italian Association for Cancer Research (Associazione Italiana per la Ricerca sul Cancro, AIRC IG16977).

Acknowledgments: We thank Hannu Myllykallio for proving us the expression plasmid hTS-pQE80L. We acknowledge the European Synchrotron Radiation Facility (ESRF, Grenoble, France) for providing access to the beamline ID30-B. Furthermore, we thank all the staff of the synchrotron source for their assistance in using the beamlines.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Kholodar, S.A.; Kohen, A. Noncovalent Intermediate of Thymidylate Synthase: Fact or Fiction? *J. Am. Chem. Soc.* **2016**, *138*, 8056–8059. [CrossRef] [PubMed]
2. Houghton, J.A.; Houghton, P.J. Cellular responses to antimetabolite anticancer agents: Cytostasis versus cytotoxicity. In *Progress in Cell Cycle Research*; Springer: Boston, MA, USA, 1996; pp. 175–185. ISBN 978-1-4613-7693-4.
3. Taddia, L.; D'Arca, D.; Ferrari, S.; Marraccini, C.; Severi, L.; Ponterini, G.; Assaraf, Y.G.; Marverti, G.; Costi, M.P. Inside the biochemical pathways of thymidylate synthase perturbed by anticancer drugs: Novel strategies to overcome cancer chemoresistance. *Drug Resist. Updates* **2015**, *23*, 20–54. [CrossRef] [PubMed]
4. Ferrari, S.; Severi, L.; Pozzi, C.; Quotadamo, A.; Ponterini, G.; Losi, L.; Marverti, G.; Costi, M.P. Chapter Seventeen—Human Thymidylate Synthase Inhibitors Halting Ovarian Cancer Growth. In *Vitamins and Hormones*; Litwack, G., Ed.; Ovarian Cycle; Academic Press: Cambridge, MA, USA, 2018; Volume 107, pp. 473–513.
5. Liu, J.; Schmitz, J.C.; Lin, X.; Tai, N.; Yan, W.; Farrell, M.; Bailly, M.; Chen, T.; Chu, E. Thymidylate synthase as a translational regulator of cellular gene expression. *Biochim. Biophys. Acta* **2002**, *1587*, 174–182. [CrossRef]
6. Schiffer, C.A.; Clifton, I.J.; Davisson, V.J.; Santi, D.V.; Stroud, R.M. Crystal structure of human thymidylate synthase: A structural mechanism for guiding substrates into the active site. *Biochemistry* **1995**, *34*, 16279–16287. [CrossRef]
7. Phan, J.; Koli, S.; Minor, W.; Dunlap, R.B.; Berger, S.H.; Lebioda, L. Human Thymidylate Synthase Is in the Closed Conformation When Complexed with dUMP and Raltitrexed, an Antifolate Drug. *Biochemistry* **2001**, *40*, 1897–1902. [CrossRef]
8. Chu, E.; Allegra, C.J. The role of thymidylate synthase as an RNA binding protein. *Bioessays* **1996**, *18*, 191–198. [CrossRef] [PubMed]
9. Berger, S.H.; Berger, F.G.; Lebioda, L. Effects of ligand binding and conformational switching on intracellular stability of human thymidylate synthase. *Biochim. Biophys. Acta* **2004**, *1696*, 15–22. [CrossRef]
10. Lin, X.; Liu, J.; Maley, F.; Chu, E. Role of cysteine amino acid residues on the RNA binding activity of human thymidylate synthase. *Nucleic Acids Res.* **2003**, *31*, 4882–4887. [CrossRef]
11. Voeller, D.M.; Zajac-Kaye, M.; Fisher, R.J.; Allegra, C.J. The identification of thymidylate synthase peptide domains located in the interface region that bind thymidylate synthase mRNA. *Biochem. Biophys. Res. Commun.* **2002**, *297*, 24–31. [CrossRef]
12. Salo-Ahen, O.M.H.; Tochowicz, A.; Pozzi, C.; Cardinale, D.; Ferrari, S.; Boum, Y.; Mangani, S.; Stroud, R.M.; Saxena, P.; Myllykallio, H.; et al. Hotspots in an Obligate Homodimeric Anticancer Target. Structural and Functional Effects of Interfacial Mutations in Human Thymidylate Synthase. *J. Med. Chem.* **2015**, *58*, 3572–3581. [CrossRef]

13. Chen, D.; Jansson, A.; Sim, D.; Larsson, A.; Nordlund, P. Structural analyses of human thymidylate synthase reveal a site that may control conformational switching between active and inactive states. *J. Biol. Chem.* **2017**, *292*, 13449–13458. [CrossRef]
14. Cardinale, D.; Guaitoli, G.; Tondi, D.; Luciani, R.; Henrich, S.; Salo-Ahen, O.M.H.; Ferrari, S.; Marverti, G.; Guerrieri, D.; Ligabue, A.; et al. Protein–protein interface-binding peptides inhibit the cancer therapy target human thymidylate synthase. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, E542–E549. [CrossRef]
15. Studier, F.W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **2005**, *41*, 207–234. [CrossRef]
16. Benvenuti, M.; Mangani, S. Crystallization of soluble proteins in vapor diffusion for X-ray crystallography. *Nat. Protoc.* **2007**, *2*, 1633–1651. [CrossRef] [PubMed]
17. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132. [CrossRef] [PubMed]
18. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 72–82. [CrossRef] [PubMed]
19. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.W.; McCoy, A.; et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242. [CrossRef]
20. Vagin, A.; Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 22–25. [CrossRef]
21. Murshudov, G.N.; Skubák, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355–367. [CrossRef]
22. Winn, M.D.; Isupov, M.N.; Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D Biol. Crystallogr.* **2001**, *57*, 122–133. [CrossRef]
23. Painter, J.; Merritt, E.A. TLSMD web server for the generation of multi-group TLS models. *J. Appl. Cryst.* **2006**, *39*, 109–111. [CrossRef]
24. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [CrossRef]
25. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486–501. [CrossRef]
26. Langer, G.; Cohen, S.X.; Lamzin, V.S.; Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **2008**, *3*, 1171–1179. [CrossRef]
27. Laskowski, R.A.; MacArthur, M.W.; Thornton, J.M. Validation of protein models derived from experiment. *Curr. Opin. Struct. Biol.* **1998**, *8*, 631–639. [CrossRef]
28. Potterton, L.; McNicholas, S.; Krissinel, E.; Gruber, J.; Cowtan, K.; Emsley, P.; Murshudov, G.N.; Cohen, S.; Perrakis, A.; Noble, M. Developments in the CCP4 molecular-graphics project. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2288–2294. [CrossRef]
29. Pedersen-Lane, J.; Maley, G.F.; Chu, E.; Maley, F. High-Level Expression of Human Thymidylate Synthase. *Protein Expr. Purif.* **1997**, *10*, 256–262. [CrossRef]
30. Stover, P.; Schirch, V. The metabolic role of leucovorin. *Trends Biochem. Sci.* **1993**, *18*, 102–106. [CrossRef]
31. Pozzi, C.; Ferrari, S.; Cortesi, D.; Luciani, R.; Stroud, R.M.; Catalano, A.; Costi, M.P.; Mangani, S. The structure of Enterococcus faecalis thymidylate synthase provides clues about folate bacterial metabolism. *Acta Cryst. D* **2012**, *68*, 1232–1241. [CrossRef]
32. Brunn, N.D.; Dibrov, S.M.; Kao, M.B.; Ghassemian, M.; Hermann, T. Analysis of mRNA recognition by human thymidylate synthase. *Biosci. Rep.* **2014**, *34*, e00168. [CrossRef]
33. Lin, Q.; Parker, M.J.; Taguchi, A.T.; Ravichandran, K.; Kim, A.; Kang, G.; Shao, J.; Drennan, C.L.; Stubbe, J. Glutamate 52-β at the α/β subunit interface of *Escherichia coli* class Ia ribonucleotide reductase is essential for conformational gating of radical transfer. *J. Biol. Chem.* **2017**, *292*, 9229–9239. [CrossRef]
34. Sayre, P.H.; Finer-Moore, J.S.; Fritz, T.A.; Biermann, D.; Gates, S.B.; MacKellar, W.C.; Patel, V.F.; Stroud, R.M. Multi-targeted antifolates aimed at avoiding drug resistance form covalent closed inhibitory complexes with human and *Escherichia coli* thymidylate synthases. *J. Mol. Biol.* **2001**, *313*, 813–829. [CrossRef] [PubMed]
35. Almqvist, H.; Axelsson, H.; Jafari, R.; Dan, C.; Mateus, A.; Haraldsson, M.; Larsson, A.; Molina, D.M.; Artursson, P.; Lundbäck, T.; et al. CETSA screening identifies known and novel thymidylate synthase inhibitors and slow intracellular activation of 5-fluorouracil. *Nat. Commun.* **2016**, *7*, 11040. [CrossRef]

36. Almog, R.; Waddling, C.A.; Maley, F.; Maley, G.F.; Van Roey, P. Crystal structure of a deletion mutant of human thymidylate synthase Δ(7–29) and its ternary complex with Tomudex and dUMP. *Protein Sci.* **2001**, *10*, 988–996. [CrossRef] [PubMed]
37. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774–797. [CrossRef]

Chapter 5

Conclusion

hTS is considered an important target for anticancer chemotherapy [43, 85, 46]. Indeed, various hTS inhibitors, targeting the enzyme active site, are currently used in clinic as anticancer agents. Nonetheless, these hTS-targeting drugs have serious limitations due to the onset of resistance. New strategies are thus needed to effectively block the activity of this enzyme without triggering drug-resistance phenomena in cells [43, 58]. Targeting allosteric-binding sites or the hTS homodimer interface are promising strategies; even so, the development effective interface binders is challenging. On this purpose, the introduction of point mutations at the hTS dimer interface can perturb the stability of the enzyme quaternary, helping the design and development of effective interface-perturbing drugs. In the *first section* of this work, we have investigated two hTS interface mutants in which Phe59 and Tyr202 have been singularly replaced by alanine residues. Former investigation on these variants pointed out the role of both residue as interface textithot spots in hTS, showing an effective destabilization of the enzyme quaternary assembly upon their alanine mutation. Here, we have investigated the thermal stability profiles of both hTS F59A and Y202A variants, showing a slight reduction of their T_m values with respect to the wild-type enzyme, indicative of the destabilization effects induced by the mutations. Despite the attempting made to obtain crystals of the hTS F59A interface variant, no structural information is yet available on it. On the other hand, we have solved the X-ray crystal structures of the hTS Y202A variant and of its complex with the inhibitor 1. The structure of the complex, shows that 1 binds at the periphery of the dimer interface, partially matching the binding site of the LR peptide, shown to act as allosteric inhibitor of hTS by stabilizing the inactive conformation of the enzyme [58]. The structural information achieved on this complex will guide the rational design of molecules able to effectively bind hTS outside the catalytic cavity, thus acting as enzyme inhibitors by different mechanisms of action, such as allosteric inhibition. In the second section, we have generated and investigated three hTS interface variants in which the interface residues Gln62 and Thr251 have been replaced by the bulkier charged amino acids arginine and glutamate, respectively. The interface variants Q62R, T251E, and Q62R-T251E have been characterized through CD thermal denaturation and kinetic analyses, showing that these mutations drastically affect the stability and the catalytic activity of the enzyme. Indeed, all variants showed significantly decreased T_m values and enzymatic activity with respect to HT-hTS. To deeply investigate the effects of these mutations we have also characterized the structure of all variants and compared them to the information available on the wild-type enzyme. All interface mutants show a slight aperture of the hTS dimer interface, reasonably accounting for the destabilization effect observed by thermal denaturation analysis. We also observed that these interface variants have a peculiar behavior, resulting from an altered equilibrium between the active and inactive conformations of the enzyme. Deeper investigations are needed to unveil the effective contribution of interface mutation on the active/inactive conformation equilibrium of hTS. Nonetheless, the generation of hTS variants having a more accessible interface area can facilitate the screening of interface-targeting molecules, providing relevant information for the rational design of innovative hTS inhibitors.

Chapter 6

Targeting methyltransferases in human pathogenic bacteria: insights into thymidylate synthase (TS) and flavin-dependent TS (FDTS)

6.1 Brief introduction to the review

Thymidylate synthases are pivotal enzymes also for bacterial cells where they provide the only *de novo* source of dTMP, required for DNA synthesis. In bacteria, two families of thymidylate synthases have been identified: folate-dependent and flavin-dependent thymidylate synthases (TS, EC 2.1.1.45, and FDTS, EC 2.1.1.148, respectively). Due to their fundamental role, TS and FDTS are validated targets for the development of antibiotic drugs. In this review, we describe the recent advances in the structural and functional characterization of bacterial TSs and FDTSs, the current understanding of their mechanisms of action and the recent progresses in the development of inhibitors targeting these enzymes in human pathogenic bacteria. TS and FDTS, encoded by the *thyA* and *thyX* genes, respectively, are highly divergent enzymes, characterized by different catalytic mechanisms and cofactors [46, 133]. At variance with TS that relies only on mTHF, FDTS requires mTHF, flavin adenine dinucleotide (FAD) and nicotinamide adenine dinucleotide phosphate (NADPH) to perform its action[46, 51, 133, 134]. In the TS-catalyzed reaction, mTHF provides both the methylene group and the hydride required to convert dUMP in dTMP [46, 51, 135]. On the other hand, the two additional cofactors NADPH and FAD allow FDTSs to combine the TS and DHFR functions, whereas mTHF is solely used as methyl donor to yield DHF [51, 133]. The mechanism of action of FDTSs is complicate and it has not been fully elucidated yet. Mechanistic studies have been also performed on bacterial TSs, allowing to explain its catalytic activity; nonetheless, the half-site reactivity [136] is still an open debate and recent evidences are somehow controversial. The TS dimer interface has been shown to play a major contribution in the inter-subunit communication, occurring between the two enzymes activity, but its potential as drug-targetable area is yet unexplored for bacterial enzymes. On the other hand, a wide variety of molecules have been explored as TS active site inhibitors over the years, but the high conservation of this region with human counterpart constitutes an important issue for the design of selective bTS inhibitors. Even though, some relevant steps forward have been recently reported in this field. FDTSs seems more promising as targets for development of new antibacterial drugs since they have no human counterpart enzymes. Notably, some human pathogens rely only on this enzyme for dTMP production [137, 138]. The development of FDTS-targeting molecules has led to the identification of potent inhibitors; but they resulted only poorly effective *in vivo*. Although these are promising results, more efforts are required to obtain drug candidates effectively targeting bTS and FDTS enzymes.

Publication: Pozzi C, Lopresti L, Tassone G, Mangani S. Targeting Methyltransferases in Human Pathogenic Bacteria: Insights into Thymidylate Synthase (TS) and Flavin-Dependent TS (FDTS). Molecules. 2019 Apr 25;24(8)

Contribution: My contribution to this work was writing the first draft of the review and its revision.

DOI: 10.3390/molecules24081638

molecules

MDPI

Review

Targeting Methyltransferases in Human Pathogenic Bacteria: Insights into Thymidylate Synthase (TS) and Flavin-Dependent TS (FDTS)

Cecilia Pozzi *, Ludovica Lopresti, Giusy Tassone and Stefano Mangani *

Department of Biotechnology, Chemistry and Pharmacy–Department of Excellence 2018-2020, University of Siena, via Aldo Moro 2, 53100 Siena, Italy; lopresti4@student.unisi.it (L.L.); giusy.tassone@unisi.it (G.T.)
* Correspondence: pozzi4@unisi.it (C.P.); stefano.mangani@unisi.it (S.M.);
Tel.: +39-0577-232132 (C.P.); +39-0577-234255 (S.M.)

Received: 31 March 2019; Accepted: 25 April 2019; Published: 25 April 2019

check for updates

Abstract: In cells, thymidylate synthases provide the only *de novo* source of 2′-deoxythymidine-5′-monophosphate (dTMP), required for DNA synthesis. The activity of these enzymes is pivotal for cell survival and proliferation. Two main families of thymidylate synthases have been identified in bacteria, folate-dependent thymidylate synthase (TS) and flavin-dependent TS (FDTS). TS and FDTS are highly divergent enzymes, characterized by exclusive catalytic mechanisms, involving different sets of cofactors. TS and FDTS mechanisms of action have been recently revised, providing new perspectives for the development of antibacterial drugs targeting these enzymes. Nonetheless, some catalytic details still remain elusive. For bacterial TSs, half-site reactivity is still an open debate and the recent evidences are somehow controversial. Furthermore, different behaviors have been identified among bacterial TSs, compromising the definition of common mechanisms. Moreover, the redox reaction responsible for the regeneration of reduced flavin in FDTSs is not completely clarified. This review describes the recent advances in the structural and functional characterization of bacterial TSs and FDTSs and the current understanding of their mechanisms of action. Furthermore, the recent progresses in the development of inhibitors targeting TS and FDTS in human pathogenic bacteria are summarized.

Keywords: thymidylate synthase; flavin-dependent thymidylate synthase; mechanism of action; half-site reactivity; inhibitors; selectivity

---

1. Introduction

Thymidylate synthase is a class of methyltransferase enzymes required for *de novo* 2′-deoxythymidine-5′-monophosphate (dTMP) synthesis. These enzymes catalyze the methylation of 2′-deoxyuridine-5′-monophosphate (dUMP) using N^5,N^{10}-methylentetrahydrofolate (CH_2H_4folate) as co-substrate. Thymidylate synthases are pivotal for cell survival and replication since they provide the unique biosynthetic source of dTMP, essential for DNA synthesis.

In bacteria, two main families of thymidylate synthases are known, folate-dependent thymidylate synthase (TS, EC 2.1.1.45) and flavin-dependent thymidylate synthase (FDTS, EC 2.1.1.148), encoded by *thyA* and *thyX* genes, respectively [1,2]. TS and FDTS are highly divergent at all structural levels [1,2]. These enzymes are also characterized by exclusive catalytic mechanisms that involve different sets of cofactors [1–4]. At variance with TS that relies only on CH_2H_4folate, FDTS requires CH_2H_4folate, flavin adenine dinucleotide (FAD) and nicotinamide adenine dinucleotide phosphate (NADPH) to perform its action [1–4]. In the TS-catalyzed reaction, CH_2H_4folate provides both the methylene group and the hydride required to convert dUMP in dTMP (Figure 1) [1,5]. Dihydrofolate (H_2folate), generated as byproduct of the TS reaction, is then converted to tetrahydrofolate (H_4folate) through a second enzyme, dihydrofolate reductase (DHFR, encoded by *folA* gene) (Figure 1) [5]. On the other hand, FDTSs are able to combine the TS and DHFR functions, relying on the two additional cofactors, NADPH and FAD (Figure 1) [2]. FDTSs use CH_2H_4folate solely as the methyl donor, yielding H_4folate (Figure 1) [2,4]. At a later stage, the pathways of TS and FDTS converge in the recycling of the cofactor CH_2H_4folate from H_4folate, ensured by the enzyme serine hydroxymethyltransferase [5].

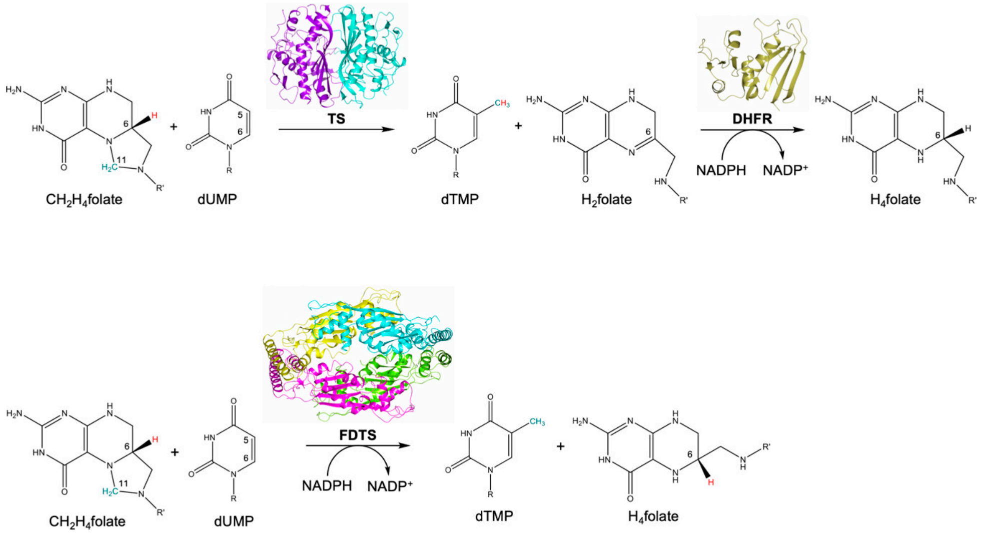


Figure 1. Reactions catalyzed by TS and DHFR (upper panel) and FDTS (lower panel) (TS, PDB id 3QJ7; DHFR, PDB id 5UIH; FDTS, PDB id 3GCW). In the FDTS catalyzed reaction, the cofactor FAD is not displayed because it is oxidized and subsequently reduced in each catalytic cycle. R = 2′-deoxyribose-5′-monophosphate; R′ = *p*-benzoyl-amino-L-glutamic acid.

At variance with TS that is present in some viruses and various organisms including humans, FDTS is unique to bacteria [1–4]. Important human pathogens, including *Helicobacter pylori*, *Borrelia Burgdorferi*, *Treponema pallidum*, *Chlamydia* species and *Rickettsia* species, rely only on FDTS for dTMP biosynthesis [2,6,7]. On the other hand, human pathogenic bacteria such as *Staphylococcus aureus*, *Enterococcus faecalis* and *Pseudomonas aeruginosa*, have only the *thyA* gene, expressing solely the TS enzyme [2,6,7]. A third group of bacteria, possessing both *thyA* and *thyX* genes, has been identified [2,6,7]. *Bacillus anthracis*, *Clostridium botulinum*, and *Mycobacterium* species are examples of important human pathogens belonging to this group [2,6,7]. In view of their common biological function, the reason concomitant expression of TS and FDTS occurs in these bacteria is not yet fully understood. Studies on *Mycobacterium tuberculosis* have evidenced that the *thyX* gene is essential, while the *thyA* deletion confers *p*-aminosalicylic acid resistance [6]. Furthermore, investigations on multi-drug resistant strains of *Mycobacterium tuberculosis*, have shown up-regulation of the *thyX* gene, responsible for FDTS overexpression [8].

Nowadays, the widespread diffusion of antibiotic resistance is an important health issue [9–12]. The major challenges are the identification of new microbial targets and the development of effective antibiotic therapies able to treat resistant infections. For this purpose, FDTS represents a promising target for the development of new antibiotics, since it has no counterpart enzyme in the human host [13,14]. On the other hand, TS is highly conserved in human and bacteria creating limitations for the development of inhibitors selectively targeting the bacterial enzyme [15]. Recent studies have provided important new insights into the catalytic process of both methyltransferase enzymes [3,4]. Indeed, new mechanisms of action for TS and FDTS have been recently proposed [3,4], opening new perspectives for the development of antibacterial drugs targeting these enzymes. This review is aimed to summarize the current understanding of structure and function of bTSs and FDTSs and the recent progresses in the development of inhibitors targeting these enzymes in human pathogenic bacteria.

2. Bacterial Thymidylate Synthases (bTSs)

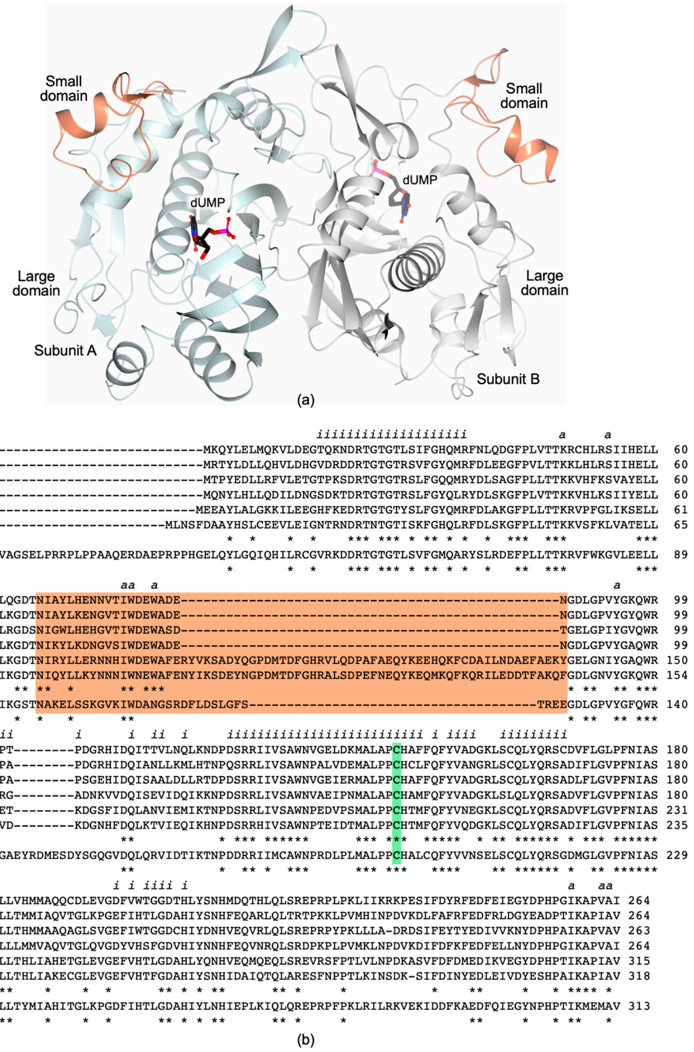
2.1. Structural Insights into bTSs from Human Pathogens

Few crystallographic structures of TSs from human pathogenic bacteria have been reported to date. The structures of TSs from *Mycobacterium tuberculosis* (*Mtb*TS; PDB id 3QJ7, unpublished research), *Staphylococcus aureus* (*Sa*TS; PDB id 4DQ1, unpublished research), *Enterococcus faecalis* (*Ef*TS; PDB id 3UWL [16]), *Escherichia coli* (*Ec*TS; PDB id 1F4B [17]), *Brucella melitensis* (*Bm*TS; PDB id 3IX6, unpublished research), and *Elizabethkingia anophelis* (*Ea*TS; PDB id 6AUJ, unpublished research) are currently available in the Protein Data Bank (PDB), but limited information is reported in the literature.

TS works as an obligate homodimer since residues from both subunits contribute to create the enzyme active site. Each subunit is composed of two domains, named large and small domains (LD and SD, respectively, Figure 2a). The LD, representing the conserved core of the enzyme, has a mixed α/β structure characterized by seven α-helices and eight β-strands (Figure 2a). A five-stranded β-sheet in the LD of each subunit generates the dimer interface, a crucial area for the enzyme function and for inter-subunit communication [18–20]. On the other hand, the SD is highly variable among bTSs, in terms of amino acid composition and length (Figures 2b and 3a). The active site is a shallow cavity embedded between the two domains, where the substrate and the cofactor bind in a sequential order, relying on dUMP binding followed by CH_2H_4folate. A recent study has evidenced that the binding order of substrate and cofactor is tightly controlled in *Ec*TS, whereas it is less stringent for the human enzyme (hTS) [21].

The substrate binding pocket is highly conserved among bTSs (Figures 2b and 3b). The dUMP phosphate moiety is bound to four conserved arginine residues (Arg21, Arg166, Arg126′ and Arg127′ in *Ec*TS), two of them protruding in the active site from the cognate subunit (Arg126′ and Arg127′ in *Ec*TS, Figure 3b). The four-arginine cluster is pivotal to anchor the substrate in its pocket, indeed substitution of these arginine residues impairs the enzyme catalytic activity [22–25]. This is also confirmed by the recent evidence of poor catalytic efficiency displayed by the TS from the non-pathogenic bacterium *Vibrio parahaemolyticus*, where Arg127′ is constitutively missing [26]. The correct orientation of the dUMP uracil ring is ensured by a conserved asparagine (Asn177 in *Ec*TS) [27]. The asparagine amide moiety forms two H-bonds with the nitrogen N3 (position 3) and the ketone oxygen in position 4, on the dUMP uracil (Figure 3b). Studies performed on *Ec*TS have revealed that the removal of the asparagine side chain disorganizes the substrate placement, drastically reducing the enzyme catalytic activity [27]. The correct orientation of the substrate is further ensured by the H-bonds between the dUMP ribose hydroxyl and conserved tyrosine and histidine residues (Tyr209 and His207 in *Ec*TS; Figure 3b). Recent evidence on *Ec*TS has shown that this conserved tyrosine is pivotal to pre-organize the active site for the hydride transfer in the last stages of the TS catalyzed reaction (the reaction mechanism is detailed in Section 2.2) [28].

The cofactor binding site is also shared among bacterial TSs (Figures 2b and 3c). The pteridine nitrogen N3 is H-bonded to the carboxylate group of a conserved aspartic residue (Asp169 in *Ec*TS; Figure 3c). Within the bTS active site, the cofactor is mainly stabilized by hydrophobic interactions, indeed it entails a network of van der Waals contacts with a set of conserved hydrophobic residues (Figures 2b and 3c). The cofactor is further H-bonded (both directly and through water mediated interactions) with the backbone of the penultimate C-terminal residue (Figure 3c). The C-terminal segment (residues 261–264 in *Ec*TS) closes on the active site after cofactor binding, sealing the cavity. The external portion of the active site is highly variable among bTSs, inducing slightly different orientations of the terminal glutamate moiety of the cofactor, which entails mainly water-mediated interactions in this area. Lys48 in *Ec*TS is the only conserved residue that characterizes this site. Lys48 forms a water-mediated interaction with the cofactor glutamate moiety, which has been proved to be important during the enzyme catalysis [29].



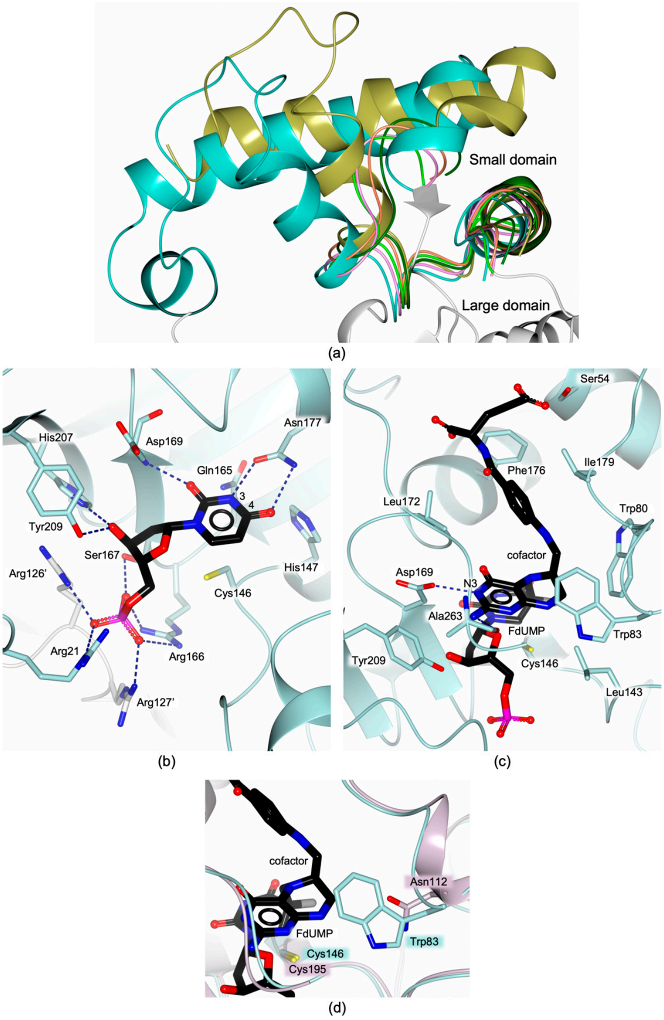
```
                                                    iiiiiiiiiiiiiiiiiiii            a     a
EcTS  1 -----------------------------MKQYLELMQKVLDEGTQKNDRTGTGTLSIFGHQMRFNLQDGFPLVTTKRCHLRSIIHELL  60
BmTS  1 -----------------------------MRTYLDLLQHVLDHGVDRDDRTGTGTRSVFGYQMRFDLEEGFPVLTTKKLHLRSIIHELL  60
MtbTS 1 -----------------------------MTPYEDLLRFVLETGTPKSDRTGTGTRSLFGQQMRYDLSAGFPLLTTKKVHFKSVAYELL  60
EaTS  1 -----------------------------MQNYLHLLQDILDNGSDKTDRTGTGTRSLFGYQLRYDLSKGFPLVTTKKVHLKSIIYELL  60
EfTS  1 ----------------------------MEEAYLALGKKILEEGHFKEDRTGTGTYSLFGYQMRFDLAKGFPLLTTKRVPFGLIKSELL  61
SaTS  1 ------------------------MLNSFDAAYHSLCEEVLEIGNTRNDRTNTGTISKFGHQLRFDLSKGFPLLTTKKVSFKLVATELL  65
                                        *  *    *  *    *** *** * ** * *  *  ***  ***         ***

hTS   1 MPVAGSELPRRPLPPAAQERDAEPRPPHGELQYLGQIQHILRCGVRKDDRTGTGTLSVFGMQARYSLRDEFPLLTTKRVFWKGVLEELL  89
                                        *       *  *    *** *** * ** * *  *   **  ***         ***

                          aa  a                                                            a
EcTS    WFLQGDTNIAYLHENNVTIWDEWADE--------------------------------------------------NGDLGPVYGKQWR  99
BmTS    WFLKGDTNIAYLKENGVTIWDEWADE--------------------------------------------------NGDLGPVYGYQWR  99
MtbTS   WFLRGDSNIGWLHEHGVTIWDEWASD--------------------------------------------------TGELGPIYGVQWR  99
EaTS    WFLKGDTNIKYLKDNGVSIWDEWADE--------------------------------------------------NGDLGPVYGAQWR  99
EfTS    WFLKGDTNIRYLLERNNHIWDEWAFERYVKSADYQGPDMTDFGHRVLQDPAFAEQYKEEHQKFCDAILNDAEFAEKYGELGNIYGAQWR 150
SaTS    WFIKGDTNIQYLLKYNNNIWNEWAFENYIKSDEYNGPDMTDFGHRALSDPEFNEQYKEQMKQFKQRILEDDTFAKQFGDLGNVYGKQWR 154
        **  ** **  *      ** ***                                                     * **  ** ***
hTS     WFIKGSTNAKELSSKGVKIWDANGSRDFLDSLGFS--------------------------------------TREEGDLGPVYGFQWR 140
        **  *  *   *      **                                                         * **  ** ***
         iii        i      i  ii  i     iiiiiiiiiiiiiiiiiiiiiiiiii i iiii   iiiiiiiii
EcTS    AWPT--------PDGRHIDQITTVLNQLKNDPDSRRIIVSAWNVGELDKMALAPCHAFFQFYVADGKLSCQLYQRSCDVFLGLPFNIAS 180
BmTS    SWPA--------PDGRHIDQIANLLKMLHTNPQSRRLIVSAWNPALVDEMALPPCHCLFQFYVANGRLSCQLYQRSADIFLGVPFNIAS 180
MtbTS   SWPA--------PSGEHIDQISAALDLLRTDPDSRRIIVSAWNVGEIERMALPPCHAFFQFYVADGRLSCQLYQRSADLFLGVPFNIAS 180
EaTS    SWRG--------ADNKVVDQISEVIDQIKKNPDSRRLIVSAWNVAEIPNMALAPCHAMFQFYVADGKLSLQLYQRSADVFLGVPFNIAS 180
EfTS    HWET--------KDGSFIDQLANVIEMIKTNPDSRRLIVSAWNPEDVPSMALPPCHTMFQFYVNEGKLSCQLYQRSADVFLGVPFNIAS 231
SaTS    DWVD--------KDGNHFDQLKTVIEQIKHNPDSRRHIVSAWNPTEIDTMALPPCHTMFQFYVQDGKLSCQLYQRSADIFLGVPFNIAS 235
         *                **           * *** ******      *** ***  *****  * ** ****** * *** ******
hTS     HFGAEYRDMESDYSGQGVDQLQRVIDTIKTNPDDRRIIMCAWNPRDLPLMALPPCHALCQFYVVNSELSCQLYQRSGDMGLGVPFNIAS 229
                          **           *  ** *  ***      *** ***   ****    ** ****** *  ** ******
                         i i iiii i                                                  a   aa
EcTS    YALLVHMMAQQCDLEVGDFVWTGGDTHLYSNHMDQTHLQLSREPRPLPKLIIKRKPESIFDYRFEDFEIEGYDPHPGIKAPVAI 264
BmTS    YALLTMMIAQVTGLKPGEFIHTLGDAHIYSNHFEQARLQLTRTPKKLPVMHINPDVKDLFAFRFEDFRLDGYEADPTIKAPIAV 264
MtbTS   YALLTHMMAAQAGLSVGEFIWTGGDCHIYDNHVEQVRLQLSREPRPYPKLLLA-DRDSIFEYTYEDIVVKNYDPHPAIKAPVAV 263
EaTS    YALLLMMVAQVTGLQVGDYVHSFGDVHIYNNHFEQVNRQLSRDPKPLPVMKLNPDVKDIFDFKFEDFELLNYDPHPGIKAPVAI 264
EfTS    YALLTHLIAHETGLEVGEFVHTLGDAHLYQNHVEQMQEQLSREVRSFPTLVLNPDKASVFDFDMEDIKVEGYDPHPTIKAPIAV 315
SaTS    YALLTHLIAKECGLEVGEFVHTFGDAHIYSNHIDAIQTQLARESFNPPTLKINSDK-SIFDINYEDLEIVDYESHPAIKAPIAV 318
        ****    *    *  *      ** * * **      ** *     *           *    **     *   * **** *
hTS     YALLTYMIAHITGLKPGDFIHTLGDAHIYLNHIEPLKIQLQREPRPFPKLRILRKVEKIDDFKAEDFQIEGYNPHPTIKMEMAV 313
        ****    *    *  *      ** * * **      ** *     *                **     *   * **   *
```

(b)

Figure 2. (**a**) Cartoon representation of the TS homodimer (*Mtb*TS, PDB id 3QJ7, unpublished research). The LD of Subunits A and B are colored light cyan and white, respectively. The SD (residues 67–87, in *Mtb*TS) is colored coral in both subunits. The position of the active site is indicated by the presence of the substrate dUMP (in sticks, black carbons). (**b**) Sequence alignment of TSs from pathogenic bacteria and the human enzyme (hTS). *Enterococcus faecalis* TS (*Ef*TS, UniProtKB: Q834R3), *Staphylococcus aureus* TS (*Sa*TS, UniProtKB: P67046), *Brucella melitensis* TS (*Bm*TS, UniProtKB: P67042), *Escherichia coli* TS (*Ec*TS, UniProtKB: P0A884), *Mycobacterium tuberculosis* TS (*Mtb*TS, UniProtKB: P9WFR9), and *Elizabethkingia anophelis* TS (*Ea*TS, UniProtKB: A0A077EAN3); human TS (hTS, UniProtKB: P04818). The present alignment includes only the TS sequences from human pathogenic bacteria for which the structure has been characterized. The SD is highlighted in coral and the catalytic cysteine in green. Conserved residues are indicated by asterisks. Residues forming the dimer interface and the active site are indicated by "*i*" and "*a*", respectively.



(a)

(b) (c)

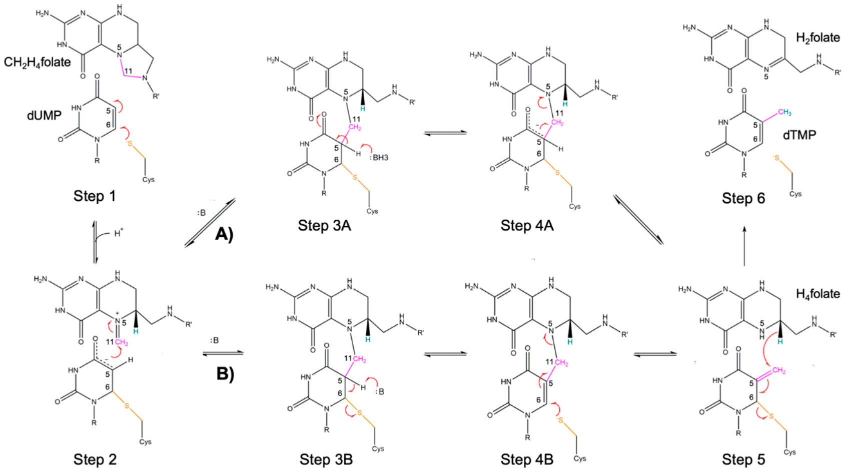
(d)

Figure 3. (**a**) Superimposition among the SD of *Mtb*TS (PDB id 3QJ7, unpublished research, coral trace), *Sa*TS (PDB id 4DQ1, unpublished research, gold trace), *Ef*TS (PDB id 3UWL [16], cyan trace), *Ec*TS (PDB id 1F4B [17], green trace), *Bm*TS (PDB id 3IX6, unpublished research, dark green trace), and *Ea*TS (PDB id 6AUJ, unpublished research, pink trace). (**b**) Conserved dUMP interactions within the substrate binding pocket of the TS active site (*Ec*TS sequence numbering). (**c**) Cofactor interactions within the active site (*Ec*TS sequence numbering). Only direct interactions are shown in Panels b and c for clarity. (**d**) Active site view of the superimposition of bTS and hTS, evidencing the presence of an unshared residue in the cofactor site, namely Asn112 in hTS that replaces the conserved tryptophan residue present in bTSs (Trp83 in *Ec*TS). In all figures, oxygen atoms are colored red, nitrogen blue, sulfur yellow, phosphorous magenta and halogen grey.

The residues responsible for substrate binding are highly conserved also in the human enzyme (Figure 2b). However, the set of hydrophobic residues composing the cofactor site is altered in hTS by the presence of Asn112 that replaces a tryptophan residue (Trp83 in *Ec*TS) shared by bTSs (Figures 2b and 3d). The presence of unshared residues in the active site of bacterial and human enzymes is of key importance for the development of selective bTS inhibitors.

2.2. Latest Updates on the TS Catalytic Mechanism

TS catalyzes the reductive methylation of dUMP to dTMP, using C_2H_4folate as co-substrate (Scheme 1) [1,30]. The catalytic cysteine (Cys146 in *Ec*TS; Figure 2b) is fundamental to activate the substrate and to attract the methylene moiety according to an established mechanism (Steps 1 and 2, Scheme 1) [3,31]. The thiol of the catalytic cysteine attacks the uracil C6 of dUMP leading to the formation of a covalent adduct (Step 1) [3,31]. The uracil C5 is thus activated to attack the imine carbon C11 of the cofactor (resulting from the aperture of the 5-membered ring of C_2H_4folate) forming the ternary adduct (Step 2). The mechanism of methyl and hydride transfer from the cofactor has been recently revised on the basis of new computational and experimental evidences [3,32–36]. According to the traditional mechanism (Path A), the transfer of the proton from the uracil C5 generates a covalent enolate intermediate (Step 3A). This stage is followed by the release of H_4folate (through Hoffman elimination, Step 4A). On the other hand, in the new mechanism (path B) the proton abstraction leads to the cleavage of the covalent bond anchoring the uracil C6 to the cysteine thiol (Step 3B). Subsequently, the intermediate undergoes 1–3 SN_2 reaction, leading to the regeneration of the covalent bond to the catalytic cysteine (Step 4B). The two mechanisms converge in the formation of the exocyclic methylene intermediate, that undergoes a concerted hydride transfer and cleavage of the C-S bond (1–3 SN_2 reaction, Step 5), preceding the formation of the products (Step 6).



Scheme 1. Proposed mechanism for the TS catalyzed reaction. The traditional (**A**) and the recently revised (**B**) paths for the central part of the reaction are displayed. R = 2′-deoxyribose-5′-monophosphate; R′ = *p*-benzoyl-amino-L-glutamic acid. Adapted with permission from Kholodar, SA & Kohen A. Noncovalent Intermediate of Thymidylate Synthase: Fact or Fiction? *J. Am. Chem. Soc.* **2016**, *138*, 8056–8059 [3]. Copyright 2016 American Chemical Society.

2.3. Asymmetry and "Half-Site Reactivity" in bTSs, a Still Open Debate

Several multi-subunit enzymes show the phenomenon of "half-site reactivity". Strictly speaking, half-site reactivity occurs when the reaction with a substrate shows a stoichiometry equal to one-half the number of identical subunits in the multimer. Half-site reactivity represents a form of negative cooperativity between the protomers, in which binding of the substrates in one catalytic site prevents catalysis in the partner subunit. Among bTSs, *Ec*TS is the most widely studied, whereas very little is known about TSs from other pathogenic bacteria. The presence of half-site reactivity in *Ec*TS, was suggested by kinetic studies, differential scanning and isothermal titration calorimetry and fluorescence quenching experiments [29,30,37,38]. Nonetheless, recent studies on *Ec*TS have provided evidence of minimal or absent negative cooperativity in substrate and cofactor binding to both catalytic sites [18,39].

Isothermal titration calorimetry (ITC) experiments have shown that the binding of substrate and cofactor occurs in both enzyme subunits with essentially no cooperativity (only a minor effect is observed in a temperature-dependent manner) [18]. On the other hand, NMR studies have evidenced a tiny, but detectable, extent of negative cooperativity for substrate and cofactor binding in the two *Ec*TS protomers [18,39]. On the basis of these investigations, it has been proposed that *Ec*TS is characterized by a "silent" or "isoergonic" allostery towards dUMP and cofactor binding (meaning that the two binding events have similar affinities but differ in ΔH, ΔS, and/or ΔC_p) [18,39]. NMR studies have also highlighted a significant contribution of the TS dimer interface in inter-subunit communication [39,40]. Meaningful chemical shift perturbations are observed across the dimer upon the population of both dUMP sites, whereas only local changes are reported when dUMP binds to only one site. The widespread perturbations induced by the second dUMP binding event indicates inter-subunit communication between the two substrate sites (and the two catalytic cavities). These perturbations involve the interface area covering the segment 146–153, Ala132, and residues 198–199 [40]. Therefore, any cooperativity is likely due to long-range, cross-interface effects, occurring upon binding of dUMP in the second site once the first is already populated [39,40]. Very recently, new evidence has been reported on *Ec*TS, upon investigating the *C*-terminal deletion variant I264Am (lacking the *C*-terminal Ile264) [19]. In TS enzymes, the *C*-terminal segment (residues 261–264 in *Ec*TS) rearranges, following the formation of the ternary adduct with substrate and cofactor, closing the cavity and contributing to stabilize the active site during catalysis. The structure of *Ec*TS I264Am shows dimer asymmetry due to different configurations of the catalytic cavities both populated by dUMP and the cofactor analogue inhibitor CB3717 (PDB id 6CDZ [19]; Figure 4a). Substrate and inhibitor are correctly aligned to mimic a catalytically-competent configuration only in the active site of one protomer (Figure 4a,b). On the other hand, the two compounds result slightly displaced in the partner subunit adopting a non-catalytically-competent configuration (Figure 4a,b). This observation supports half-site reactivity for *Ec*TS (Figure 4a). In the structure of *Ec*TS I264Am, changes are observed in the interface area, suggesting that they mediate the inter-subunit communication and regulate half-site reactivity. The most evident change concerns Phe149, facing itself on the partner subunit, for which increased sidechain flexibility is observed in *Ec*TS I264Am asymmetric dimers [19]. All together these evidences suggest that in *Ec*TS both active sites can be concomitantly populated by substrate and cofactor, but the extent of negative cooperativity at the basis of the half-site reactivity is still controversial [18,19,39,40]. However, all studies agree on the prominent role played by the dimer interface on inter-subunit communication. NMR and X-ray crystallographic studies converge on the involvement of the interface segment extending from the catalytic Cys146 to Tyr153 in inter-subunit communication [19,40]. The existence of analogous mechanisms in TSs of other pathogenic bacteria is still unknown.

In bTSs, the SD is highly variable in extension and structure, indeed *Ef*TS and *Sa*TS are characterized by a more extended SD, including an additional segment of 50 amino acids (Figure 2b). This can induce significant differences in asymmetric ligand binding and half-site reactivity among bTSs. The structural characterization of *Ef*TS has highlighted huge differences between the enzyme protomers, resulting in highly asymmetric enzyme homodimers (PDB id 3UWL [16]; Figure 4c). One subunit adopts the so-called closed conformation, in which the catalytic cavity is fully formed and the SD is structured, whereas the partner protomer is in the open form, showing a widely unstructured active site and SD (Figure 4c) [16]. Despite the structure of *Ef*TS suggests that it could be a half-site reactive enzyme, its cooperativity profile has still to be fully elucidated. Recent structural investigations have shown that substrate binding in the open-subunit of *Ef*TS induces the closure of the dUMP site, whereas the rest of the catalytic cavity keeps the open conformation (PDB id 6QYA; Figure 4d) [41]. Therefore, it has been proposed that the open/closed transition occurs in *Ef*TS as biphasic process in which dUMP binding triggers the closure of the substrate site whereas cofactor binding is required to fully structure the catalytic cavity (Figure 4d). In the *Ef*TS-dUMP complex asymmetric substrate binding is consistently observed. On the other hand, the structural characterization of *Sa*TS in complex with dUMP (*Sa*TS; PDB id 4DQ1, unpublished research) shows symmetric ligand binding, having both active sites populated by the substrate. Despite *Ef*TS and *Sa*TS have SDs of similar extension, they display different behaviors, suggesting that it is not possible to delineate common mechanisms for bTSs. Therefore, the recent evidences on *Ec*TS half-site reactivity and inter-subunit communication cannot be directly extended to all bacterial enzymes. The issues of asymmetric ligand binding and half-site reactivity in bTSs of other pathogenic bacteria remain largely unknown and are worthy of careful investigation.

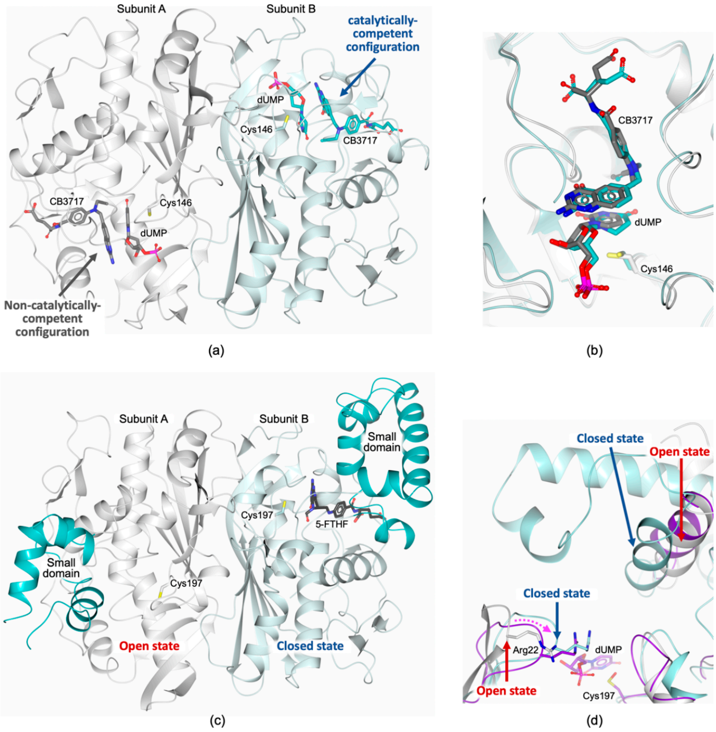


Figure 4. (**a**) Cartoon representation of *Ec*TS I264Am homodimer (Subunits A and B are colored white and light turquoise, respectively; dUMP and CB3717 are shown in sticks, color coded according to the subunit; PDB id 6CDZ [19]). In the active site of Subunit B, the substrate and the cofactor-like inhibitor CB3717 are correctly aligned to mimic a catalytically-competent configuration. On the contrary, in the partner protomer the two molecules assume a non-catalytically-competent configuration. The induced asymmetry, generated in the *Ec*TS dimer, supports the half-site reactivity of the enzyme. (**b**) Active site view of the superimposition between the catalytically-competent configuration and the non-catalytically-competent configuration of *Ec*TS I264Am homodimer. (**c**) Cartoon representation of the *Ef*TS homodimer (Subunits A and B are colored white and light turquoise, respectively, the SD is colored turquoise in both subunits; PDB id 3UWL [16]). The *Ef*TS homodimer is highly asymmetric. Subunit B is in the closed conformation, having the catalytic cavity fully formed, whereas the partner protomer is in the open form, showing a widely unstructured active site. The cofactor analogue 5-formyl-tetrahydrofolate, 5-FTHF (in sticks, black carbons), is entrapped in the active site of Subunit B (**d**) Active site view of the superimposition between the open (white cartoon and carbons) and the closed (light turquoise cartoon and carbons) states of *Ef*TS (PDB id 3UWL [16]) and the *Ef*TS–dUMP complex (purple cartoon and carbons; Cys197 is modified as S-oxycysteine; PDB id 6QYA [41]). dUMP binding in open-state subunits induces the closure of the substrate site (indicated by a dashed pink arrow), whereas cofactor binding is required to fully structure the catalytic cavity. The movement of loop 18–26, determining the closure of the substrate site, is evidenced by the shift of Arg22 (in sticks).

3. Flavin-Dependent Thymidylate Synthases (FDTSs)

3.1. Structural Insights into FDTSs from Human Pathogens

The structural characterization of FDTSs from human pathogenic bacteria is limited to the enzymes from *Mycobacterium tuberculosis* (*Mtb*FDTS; PDB id 2AF6 [42]) and *Helicobacter pylori* (*Hp*FDTS; PDB id 3N3Y [43]). FDTS are homotetrameric proteins, with four identical subunits forming extensive interactions (Figure 5a). Each subunit shows a complex fold, characterized by a central α/β domain flanked by two helical domains, known as top and bottom domains. The central domain consists of a four-stranded antiparallel β-sheet, flanked on one side by three α-helices. The four active sites are located at the interface between three neighboring monomers, having residues from all three subunits that contributes to create the catalytic cavity (Figure 5a; subunits contributing to each active site are illustrated in the figure; Active site 1, composed of residues belonging to Subunits A, B and C, is used for the structural description) [42–44]. FAD acquisition by FDTS was reported to occur inside cells, during the expression of the protein. Therefore, the purified samples already included the cofactor in its oxidized form as indicated by the peculiar yellow colorization of the protein solutions [42,43]. Within each subunit, FAD adopts an extended conformation, covering a total surface area of ~490 Å^2 [42,43]. The AMP adenine moiety of FAD is buried inside a deep pocket, whereas the AMP ribose faces the ribose of the FAD bound to a neighboring subunit (ribose of FAD^A and FAD^B, belonging to Subunits A and B, respectively, in Figure 5b; FAD^A is considered for the present description). The phosphate group of AMP and the ribityl and phosphate moieties of flavinmononucleotide (FMN), belonging to FAD^A, form H-bonds with the residues of the highly conserved FDTS motif RHR of Subunit A (residues 95^A–97^A in *Mtb*FDTS, Figure 5b; corresponding to 74^A–76^A in *Hp*FDTS). These moieties are also H-bonded with two further conserved residues belonging to Subunit B (Ser71^B and Arg190^B in *Mtb*FDTS, Figure 5b; corresponding to Ser50^B and Arg165^B in *Hp*FDTS). Inside the active site, the FAD^A isoalloxazine ring is H-bonded with the first arginine of the conserved RHR motif (Arg95^A in *Mtb*FDTS, Figure 5b; corresponding to Arg74^A in *Hp*FDTS) and with the backbone of a residue belonging to the C Subunit (Gln103^C in *Mtb*FDTS and Val82^C in *Hp*FDTS; interaction not shown in Figure 5b). The flavin ring of FAD^A is stacked with the pyrimidine of dUMP^C. The complex of *Mtb*FDTS with the substrate analogue FdUMP (PDB id 3GWC [44]) is shown in Figure 5b. FdUMP retains the same binding mode of dUMP in the *Hp*FDTS-dUMP complex (PDB id 3N3Y [43]). Notably, the dUMP^C C5 is aligned to the flavin N5 of FAD^A, separated by a short interatomic distance (<3.5 Å). The alignment of these two atoms is fundamental during the FDTS catalyzed reaction. According to the revised reaction mechanism (detailed in Section 3.2), the methyl is transferred from CH_2H_4folate to the FAD flavin N5, that subsequently donates it to the dUMP C5 allowing the formation of the product dTMP. In the catalytic cavity, the dUMP substrate (dUMP^C) is further stabilized by the interactions with a set of conserved residues (Arg95^A and Arg199^A, Arg87^C, Ser105^C and Arg107^C in *Mtb*FDTS, Figure 5b; corresponding to Arg74^A, Arg174^A, Arg66^C, Ser84^C and Arg86^C in *Hp*FDTS). The dUMP^C uracil C5 and the FAD^A flavin N5 are shielded from the solvent by a shared tyrosine (Tyr108^C and Tyr87^C in *Mtb*FDTS and *Hp*FDTS, respectively), whose phenyl moiety forms a "lid" over the reactive positions involved in the methyl transfer (Figure 5b).

The structure of *Mtb*FDTS was reported also in complex with the second cofactor NADP$^+$ that occupies the same site of FAD and adopts an analogous binding mode (PDB id 2GQ2 [45]; Figure 5c). The nicotinamide of NADP$^+$ is accommodated inside the catalytic cavity in which it replaces the isoalloxazine system of FAD (Figure 5c) [45]. It is worth noting that attempting to obtain the quaternary complex *Mtb*FDTS–FAD–BrdUMP–NADP$^+$ by co-crystallization of the enzyme with both cofactors and the substrate analogue BrdUMP, provided the structure of the binary complex *Mtb*FDTS–NADP$^+$ [45]. The mechanistic explanation of this behavior is not obvious, since it has been postulated that the regeneration of the reduced flavin during the FDTS catalyzed reaction is mediated by NADPH. Therefore, they should both bind to the enzyme during this stage. The structural characterization of FDTS in complex with both FAD and NADPH can provide key information on the mechanism of reduced flavin regeneration which has still to be fully elucidated.

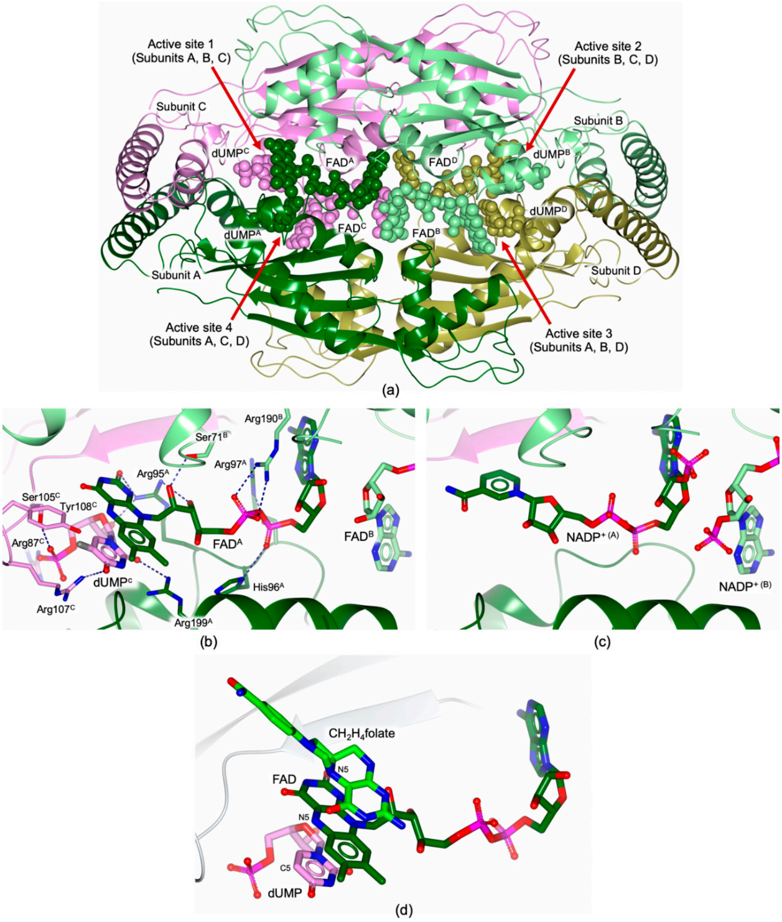
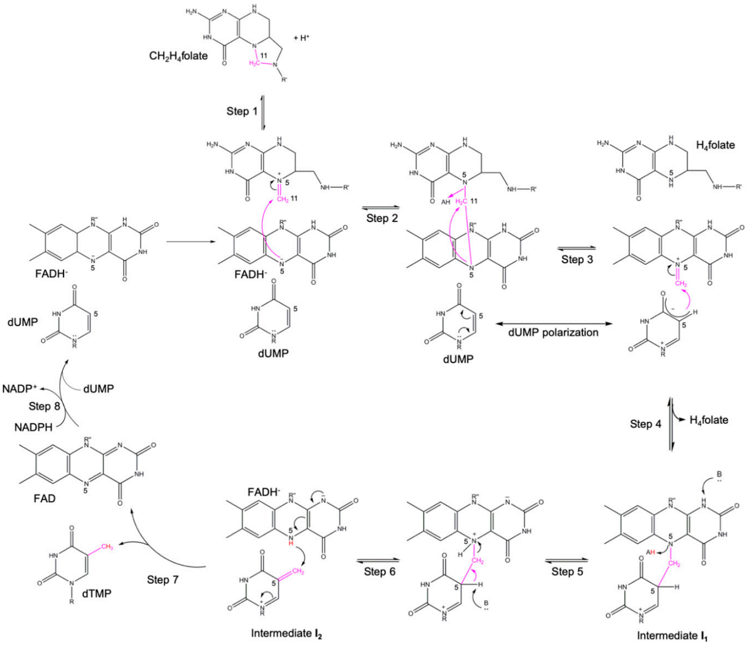


Figure 5. (**a**) Cartoon representation of the FDTS homotetramer (*Mtb*FDTS; Subunit A is colored dark green, B light green, C pink, and D gold; PDB id 3GWC [44]). Subunits contributing to each active site (numbered from 1 to 4) are illustrated in the figure. The Active site 1, composed of residues belonging to Subunits A, B and C, is used for structural description. FAD and the substrate analogue FdUMP are represented in spheres (of arbitrary radius, carbon atoms are color-coded according to the parent subunit). (**b**) H-bonds (blue dashed lines) of FAD and FdUMP (in sticks, carbon atoms are color-coded according to the parent subunit) with residues shared between *Mtb*FDTS and *Hp*FDTS (the *Mtb*FDTS sequence numbering is reported in the picture; PDB id 3GWC [44]). The Active site 1, composed of residues belonging to Subunits A, B and C, is used for structural description in Panels b and c. (**c**) NADP$^+$ (in sticks, carbon atoms are color-coded according to the parent subunit) binding mode in the complex with *Mtb*FDTS (PDB id 2GQ2 [45]). (**d**) Active site view of the quaternary complex *Tm*FDTS (white cartoon)–FAD (in sticks, dark green carbons)–dUMP (in sticks, pink carbons)–CH_2H_4folate (in sticks, green carbons) (PDB id 4GT9 [46]). The isoalloxazine moiety of FAD is sandwiched between the dUMP uracil and the pteridine of CH_2H_4folate. The pteridine N5 of CH_2H_4folate is perfectly aligned with both the flavin N5 of FAD and the uracil C5 of dUMP, configurating the active site to allow the FAD-mediated methyl transfer from CH_2H_4folate to dUMP (according to the revised reaction mechanisms described in Section 3.2). To facilitate the comparison, Panels b, c, and d have been generated using the same orientation of the FDTS active site.

Structural information concerning the CH_2H_4folate binding mode in *Mtb*FDTS and *Hp*FDTS is not available. However, the FDTS-CH_2H_4folate complex has been determined for *Thermotoga maritima* FDTS (*Tm*FDTS) [46]. The structure of the quaternary complex *Tm*FDTS–FAD–dUMP–CH_2H_4folate (PDB id 4GT9 [46]) shows the isoalloxazine moiety of FAD sandwiched between the dUMP uracil and the pteridine of CH_2H_4folate (Figure 5d). In the structure, the pteridine N5 of CH_2H_4folate is perfectly aligned with both the flavin N5 of FAD and the uracil C5 of dUMP. This configuration of the active site allows the FAD-mediated methyl transfer from CH_2H_4folate to dUMP in accord with the revised reaction mechanism described in Section 3.2.

3.2. FDTS Reaction Mechanism

The combined results of mass spectrometry, NMR and X-ray crystallographic studies on FDTSs have led Kohen and coworkers to formulate a new reaction mechanism for this class of enzymes, reported in Scheme 2 [4]. The proposed multistep process for FDTSs starts with the protonation of C_2H_4folate (Step 1), leading to the formation of a Schiff base, that is activated to transfer the methyl moiety to the flavin N5 of the reduced cofactor FADH (Step 2). In the resulting covalent adduct, the protonation of the folate N5 is mediated by an unknown species (Step 3). Once the methyl moiety is transferred on the FAD cofactor, a new Schiff base is formed. The Schiff base can react with the polarized dUMP (Step 4) to form a covalent intermediate (Intermediate I_1) in which the methyl group bridges the FADH N5 and the dUMP C5. The generation of the intermediate I_1 is accompanied by the release of H_4folate from its site (Step 4). The intermediate I_1 undergoes protonation on the flavin N5 by an unknown species (Step 5) which precedes the abstraction of the proton from the dUMP C5, still by an uncharacterized actor (Step 6). The deprotonation of the C5 induces the transfer of the methylene group on dUMP (Step 6), leading to the formation of the Intermediate I_2. I_2 undergoes a redox reaction in which flavin is oxidized at the expenses of the nucleotide that is reduced to dTMP (Step 7) and subsequently released from its site. The reduced flavin is regenerated by the oxidation of NADPH to $NADP^+$ in a second redox reaction (Step 8).



Scheme 2. Recently proposed reaction mechanism for FDTSs [4]. R = 2′-deoxyribose-5′-monophosphate; R′ = *p*-benzoyl-amino-L-glutamic acid; R″ = adenosine-5′-pyrophosphate-ribityl. Adapted from Reference [4].

4. Inhibitors of TSs and FDTSs from Human Pathogenic Bacteria

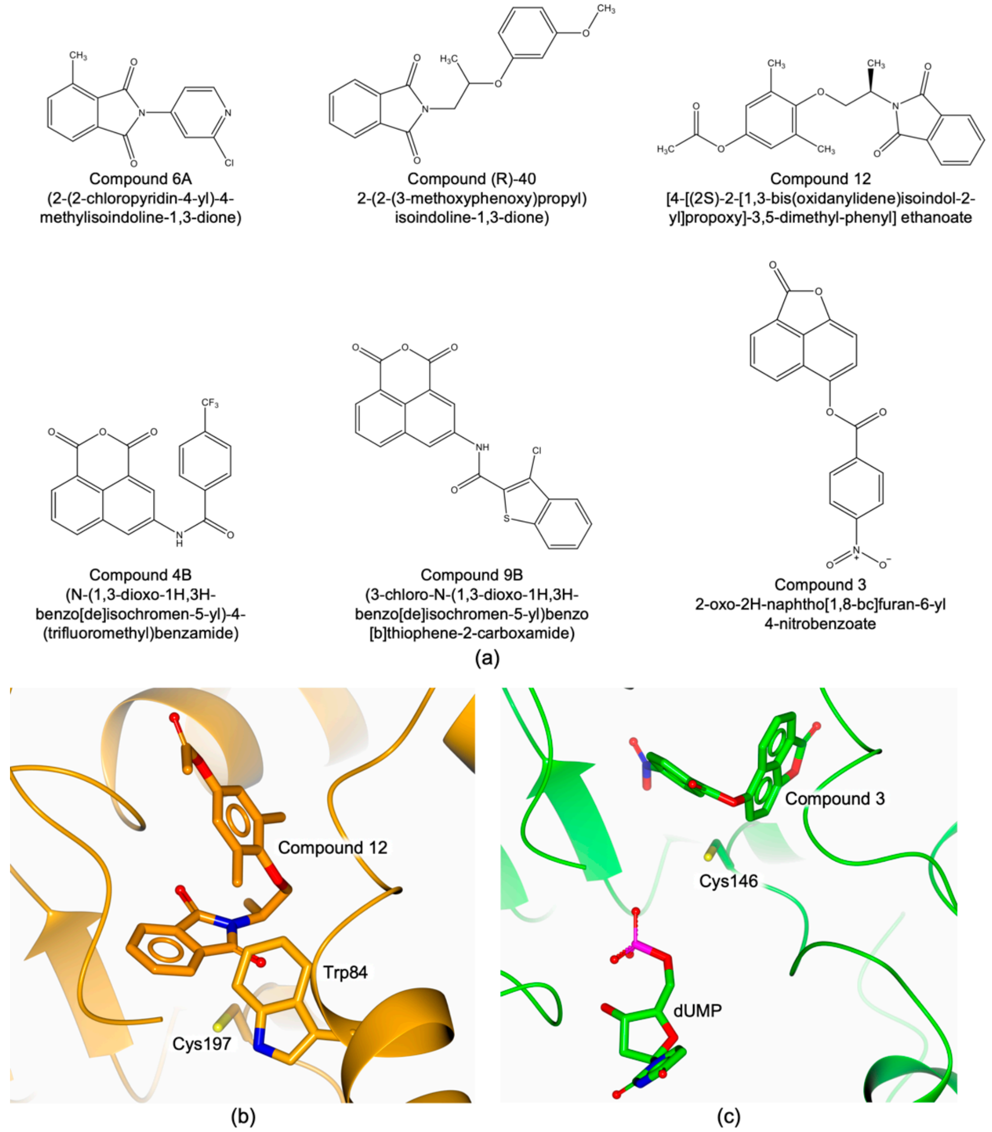
4.1. Inhibitors of Bacterial TS from Human Pathogens

The main difficulty to develop bTS inhibitors is related to selectivity. The active sites of human and bacterial TSs are highly conserved, the main difference being one unshared residue in the cofactor binding site. Furthermore, variable conformational changes have been characterized in bTSs that are different from those observed in the human enzyme. bTSs shift between open and closed conformations involving structural rearrangements that are peculiar of each bacterial enzyme [16,19,39]. However, hTS switches between the active and inactive conformations, by shifting the catalytic cysteine from the catalytic cavity to the dimer interface [47,48]. In the last decade, different molecular classes have been investigated to selectively target TSs of human pathogenic bacteria. *In silico* studies combined with structural investigations led to the identification of some phtalimide derivatives as selective bTS inhibitors [49,50]. Compounds **6A** and **(*R*)-40** (Figure 6a) resulted effective towards *Ef*TS (K_i of 7.0 µM and 13.0 µM, respectively), without affecting the human enzyme [49,50]. The structure of *Ef*TS in complex with a representative phtalimide derivative showed that Compound **12** (Figure 6a, PDB id 4O7U [50]) populates the cofactor site regardless the binding of the substrate (Figure 6b). This configuration is unusual in TS, since cofactor-like inhibitors normally populate their site in presence of dUMP (ternary complexes). The six-membered aromatic ring of phtalimide forms a close van der Waals contact with the *Ef*TS Trp84 (Figure 6b). This residue, shared among bTSs, is replaced by an asparagine (Asn112) in human TS (Figure 3d), providing basis to explain the selectivity profile of these molecules.

Phenolnaphthalein derivatives were also proposed as selective bTS inhibitors [51]. Compounds **4B** and **9B** (Figure 6a) showed relevant selectivity, being active towards *Ec*TS (K_i of 6.4 µM and 6.5 µM, respectively) and completely inactive against hTS [51]. Attempting to further improve this class of compounds through the generation of naphthofuranon derivatives resulted in an almost complete loss of selectivity, being also active towards hTS [52]. The structure of *Ec*TS obtained in complex with a representative member of this series, revealed that Compound **3** (Figure 6a, PDB id 4LRR) binds inside the catalytic cavity inducing the rotation of dUMP outside the substrate binding pocket (Figure 6c) [52].

In vitro analysis on pyrimidine-5-carbonitrile derivatives [53] and on the ruthenium-based complex $[(C_6H_6)RuL(N,N)Cl]$ [54] reported antimicrobial activity on *S. aureus* and other human pathogenic bacteria. *In silico* studies have identified them as potential *Sa*TS inhibitors [53,54]; however no experimental evidences have proven their activity towards *Sa*TS.

Recently, the structures of *Mtb*TS in complex with raltitrexed and pemetrexed (Figure 7a), two known hTS inhibitors, have been reported in the PDB (PDB ids 4FOX and 4FQS, respectively; unpublished research). The comparison between the structures *Mtb*TS–dUMP–raltitrexed and hTS–dUMP–raltitrexed (PDB id 5X5Q [48]) shows the inhibitor bicyclic system in two different orientations inside the catalytic cavities of the two enzymes (Figure 7b). However, the binding mode of pemetrexed is almost conserved in the active sites of *Mtb*TS and hTS (Figure 7c). Differences in the ligand binding mode between *Mtb*TS and hTS are exploitable for the development of selective bTS inhibitors.



Compound 6A
(2-(2-chloropyridin-4-yl)-4-methylisoindoline-1,3-dione)

Compound (R)-40
2-(2-(3-methoxyphenoxy)propyl) isoindoline-1,3-dione)

Compound 12
[4-[(2S)-2-[1,3-bis(oxidanylidene)isoindol-2-yl]propoxy]-3,5-dimethyl-phenyl] ethanoate

Compound 4B
(N-(1,3-dioxo-1H,3H-benzo[de]isochromen-5-yl)-4-(trifluoromethyl)benzamide)

Compound 9B
(3-chloro-N-(1,3-dioxo-1H,3H-benzo[de]isochromen-5-yl)benzo [b]thiophene-2-carboxamide)

Compound 3
2-oxo-2H-naphtho[1,8-bc]furan-6-yl 4-nitrobenzoate

(a)

(b) (c)

Figure 6. (**a**) Chemical structure of bTS inhibitors. (**b**) Binding mode of Compound **12** (in sticks, orange carbons) within the active site of *Ef*TS (orange cartoon and carbons) (PDB id 4O7U [50]). (**c**) Binding mode of Compound **3** (in sticks, green carbons) within the active site of *Ec*TS (green cartoon and carbons; dUMP in sticks) (PDB id 4LRR [52]).

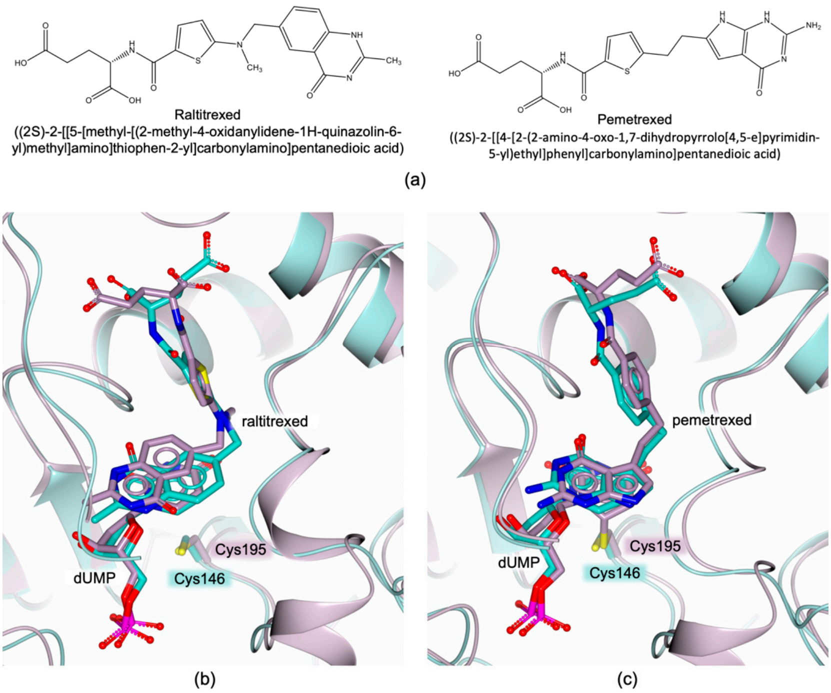


Figure 7. (**a**) Chemical structures of raltitrexed and pemetrexed. (**b**) Active site view of the superimposition of the complexes *Mtb*TS–dUMP–raltitrexed (pale turquoise cartoon and carbons; PDB id 4FOX; unpublished research) and hTS–dUMP–raltitrexed (lilac cartoon and carbons; PDB id 5X5Q [48]). (**c**) Active site view of the superimposition of the complexes *Mtb*TS–dUMP–pemetrexed (pale turquoise cartoon and carbons; PDB id 4FQS; unpublished research) and hTS–dUMP–pemetrexed (lilac cartoon and carbons; PDB id 1JUJ [55]).

4.2. FDTS Inhibitors towards Human Pathogenic Bacteria

Microbial FDTSs have no structural homology with hTS, being highly divergent in the configurations of the active sites and in their catalytic mechanisms. Therefore, FDTSs represent promising targets for the development of new antimicrobial drugs.

Natural compounds are an important source to identify new chemical scaffolds. A library of more than 2300 natural compounds was screened towards FDTSs from various pathogenic bacteria (including *Hp*FDTS, *Mtb*FDTS, and *Chlamydia trachomatis* FDTS, *Ct*FDTS) [56]. In this study, 1,4-naphthoquinone (NQ) derivatives were identified as FDTS inhibitors. Subsequent studies on the NQ derivative **C8**–**C1** (Figure 8a) showed that it is a potent inhibitor of *Hp*FDTS (estimated K_i of 367 nM), displaying meaningful antimicrobial activity on *H. pylori* (MIC 10 μg mL^{-1}) [57]. The structure of **C8-C1** in complex with the FDTS from *Paramecium bursaria chlorella* virus (*PBC*V1 FDTS; having an almost conserved active site with respect to bacterial FDTSs) showed that the compound occupies the catalytic cavity, replacing the dUMP uracil (Figure 8b; PDB id 4FZB [56]).

NQ derivatives were further developed, resulting in two compounds with improved potency towards *Hp*FDTS [57]. Compound **010-C** (Figure 8a) was the most potent *Hp*FDTS inhibitor, having a K_i in the low nano-molar range (K_i of 28 nM) [57]. These improved NQ derivatives resulted also effective towards *H. pylori* (MIC ranging from 0.625 to 10 μg mL^{-1}). The three most potent compounds of this series were also investigated *in vivo* using a mouse model for *H. pylori* infection. The compounds were tolerated in mice, but displayed a modest antibacterial effect [57].

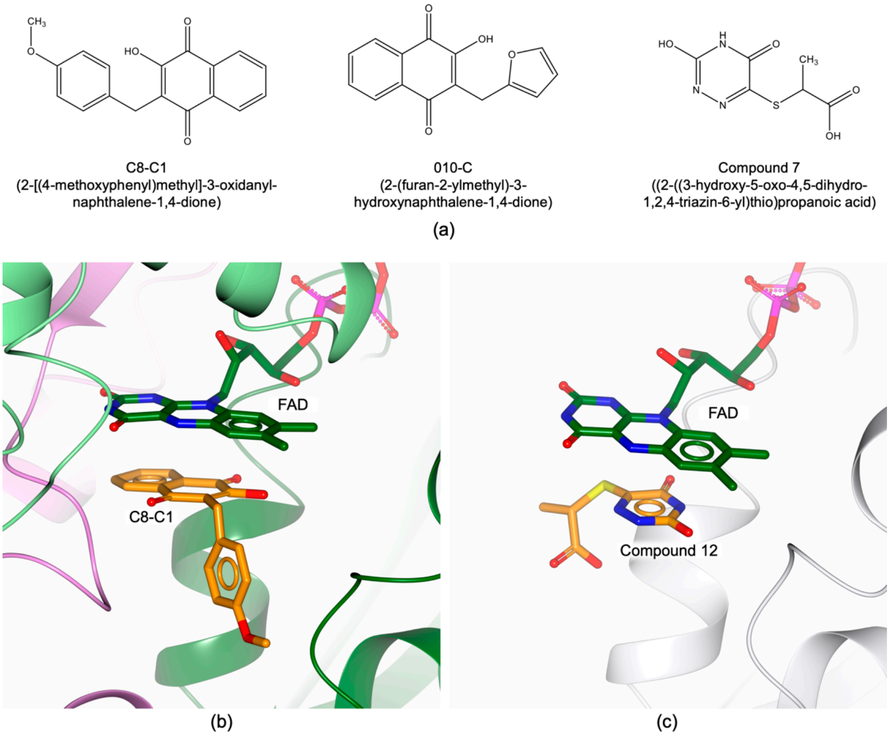


Figure 8. (**a**) Chemical structures of FDTS inhibitors. (**b**) Binding mode of the NQ derivative **C8-C1** (in sticks, orange carbons) in the active site of *PBCV*1FDTS (in cartoon, Subunit A is colored dark green, B light green, and C pink; FAD in sticks, dark green carbons; PDB id 4FZB [56]). (**c**) Binding mode of Compound **12** (in sticks, orange carbons) in the active site of *Tm*FDTS (white cartoon; FAD in sticks, dark green carbons; PDB id 5CHP [58]).

In 2016, *in silico* studies on FDTS were combined with structural investigations leading to the identification of a new molecular scaffold exploitable for the development of novel FDTS inhibitors [58]. The binding mode of Compound **7** (Figure 8a, 6% inhibition at 100 µM towards *Mtb*FDTS) was clarified in complex with the FDTS from the non-pathogenic *Thermotoga maritima* (*Tm*FDTS, PDB id 5CHP [58]). The structure shows that Compound **7** occupies the dUMP site within the catalytic cavity. Biochemical assays performed on hTS, evidenced the ineffectiveness of the compound on this target, indicating this scaffold useful for the development of novel FDTS inhibitors.

Recently, an high-throughput screening has been performed on *Mtb*FDTS, using a library of 40,000 compounds [59]. *In silico* studies have led to the identification of 1,4-benzoxazine derivatives as FDTS inhibitors. Further investigations have proven that Compound **B1-PP146** (unreported chemical structure) was the most potent 1,4-benzoxazine derivative of this series. This compound is effective on *Mtb*FDTS (IC_{50} of 0.71 µM), displaying competitive inhibition against CH_2H_4folate [59].

4.3. dUMP-Like Inhibitors towards Human Pathogenic Bacteria Expressing Both TS and FDTS Enzymes

Bacillus anthracis, *Clostridium botulinum*, and *Mycobacterium* species are examples of human pathogenic bacteria expressing both TS (*thyA* gene) and FDTS (*thyX* gene) enzymes [2]. Studies performed on *Mycobacterium tuberculosis* showed that the *thyX* gene is essential for bacteria survival and its overexpression was observed in multi-drug resistant strains [6,8]. However, the *thyA* gene was reported less essential for *M. tuberculosis* but connected with the resistance to *p*-aminosalicylic acid [6]. The effects of FDTS inhibition on bacteria expressing both methyltransferase enzymes have to be fully elucidated and the potential involvement of TS as metabolic bypass to FDTS inhibition is not excluded. The ability of these enzymes to recognize the same substrate suggested its analogue FdUMP (Figure 9a) as dual TS–FDTS inhibitor [60,61]. Indeed, FdUMP was reported as potent inhibitor of both *Mtb*TS (K_i of 2 nM) and *Mtb*FDTS (K_i 100 nM) (Figure 9b,c), exhibiting also a remarkable antimycobacterial activity (MIC 3.1 μM) [60,61].

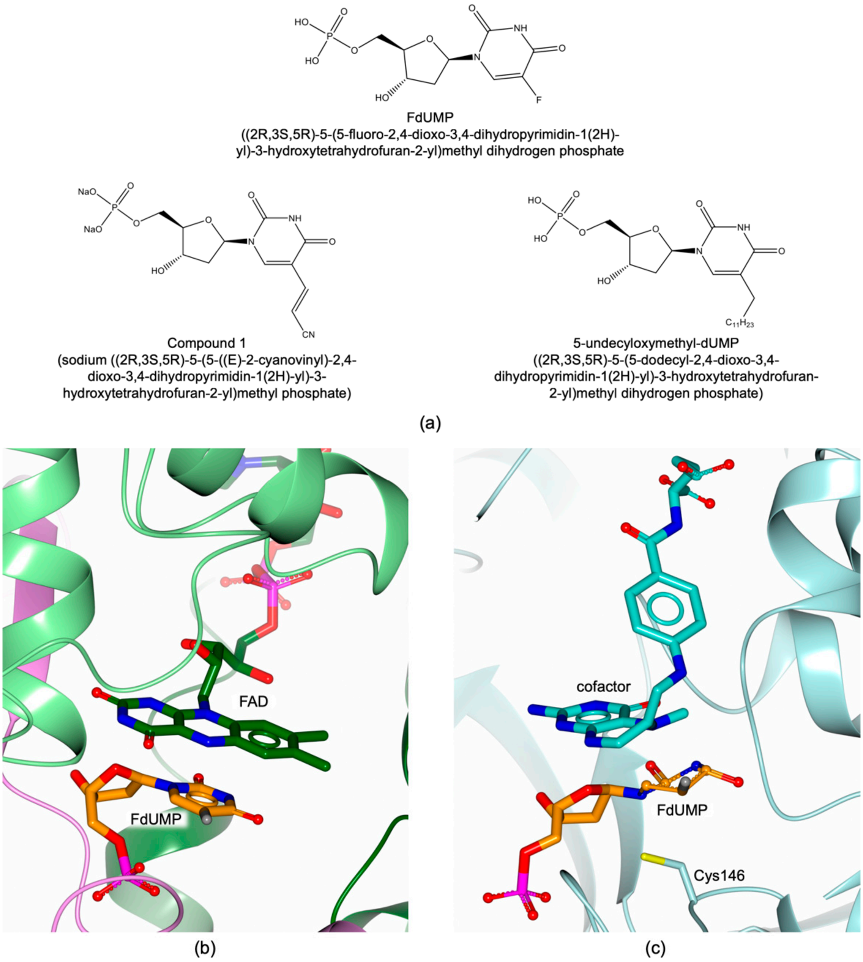


Figure 9. (**a**) Chemical structures of FdUMP and substrate analogue inhibitors. (**b**) Binding mode of FdUMP (in sticks, orange carbons) in the active site of *Mtb*FDTS (in cartoon, Subunit A is colored dark green, B light green, and C pink; FAD in sticks, dark green carbons; PDB id 3GWC [44]). (**c**) Binding mode of FdUMP (in sticks, orange carbons) in the active site of *Mtb*TS (pale turquoise cartoon and carbons; CH_2H_4folate in sticks, turquoise carbons; PDB id 4FOG; unpublished research).

In 2011, a series of 5-substituted-dUMP analogues was probed towards *Mtb*FDTS and *Mtb*TS [62]. Compound **1** (Figure 9a) resulted the most potent *Mtb*FDTS inhibitor (IC_{50} of 0.91 μM), showing no noticeable activity on *Mtb*TS ($IC_{50} > 50$ μM) [62]. The replacement of dUMP with 6-aza-dUMP in a new series of 5-substituted derivatives induced a drop of their potency towards *Mtb*FDTS (maximal inhibition of 40% at 50 μM), resulting inactive also against *Mtb*TS [63]. Analogously, the replacement of the nucleotide with an acyclic nucleoside phosphonate moiety (ANP derivatives) reduced the activity towards *Mtb*FDTS (maximal inhibition of 43% at 50 μM) [64].

Structure-activity relationship analysis combined with *in silico* studies led to the development of 5-undecyloxymethyl-dUMP (Figure 9a), for which an IC_{50} of 8.32 μM against *Mtb*FDTS was reported [65]. On the other hand, no detectable inhibition was observed on *Mtb*TS (up to concentrations of 100 μM) [65]. This compound, together with other derivatives of this series, displayed a significant antibacterial activity (MIC 10–20 μg mL^{-1}) towards two mycobacterial strains, the virulent laboratory H37Rv strain and the multi-drug resistant MS-115 [66].

Very recently, a further series of nucleotide analogues has been investigated [67]. Various 5-modified 6-aza- and 2-thio-6-aza-2′-deoxyuridine derivatives have shown a significant antibacterial activity on important human pathogens, including *Mycobacterium smegmatis*, *Staphylococcus aureus* and *Pseudomonas aeruginosa* [67]. Docking studies have predicted that they target *Mtb*FDTS. Nonetheless, their activity on *Staphylococcus aureus* and *Pseudomonas aeruginosa*, lacking the *thyX* gene, strongly suggests that these compounds inhibit other bacterial enzymes.

5. Conclusions

Methyltransferase enzymes of human pathogenic bacteria represent important targets for the development of new antibiotic drugs. The catalytic processes of TS and FDTS have been recently revised, but their mechanisms of action are not yet fully elucidated. For bTSs, half-site reactivity is still an open debate. Recent insights obtained on *Ec*TS are somehow controversial [18,19,39,40], suggesting that deeper investigations are required to unveil the extent of inter-subunit cooperativity in this enzyme. The TS dimer interface area has been shown to play a major contribution in the inter-subunit communication occurring between the two holoenzyme halves during the catalytic process. The TS dimer interface is crucial for the enzyme activity, but its potential as drug-targetable area is yet unexplored for bacterial enzymes.

FDTSs represent a promising target for development of new antibacterial drugs since they have no human counterpart enzymes. On the other hand, in FDTS, the mechanism of the redox reaction responsible for the regeneration of reduced flavin is not fully elucidated. During this process, flavin is reduced at the expenses of NADPH that is oxidized to $NADP^+$. Understanding how FAD and NADPH interact with FDTS during this stage of the catalytic reaction would represent a breakthrough to elucidate the mechanism by which the flavin cofactor is regenerated by these enzymes.

The development of FDTS-targeting molecules has led to the identification of potent inhibitors. Nonetheless, the *in vivo* properties of these compounds have to be improved to achieve suitable drug candidates. Targeting this enzyme is crucial not only in human pathogens relying only on FDTS, but also in those expressing both FDTS and TS enzymes. Although blocking the FDTS activity is pivotal, the expression of TS could provide a metabolic bypass to FDTS inhibition, leading to resistance. The importance of developing bTS inhibitors is further evidenced by the existence of important human pathogens relying only on TS. The high conservation of the active site among bacterial and human TSs constitutes an important issue for the design of selective bTS inhibitors. Relevant steps forward have been recently reported opening new perspectives for the development of effective and selective bTS-targeting drugs. The recent structural evidence acquired on bTSs from human pathogenic bacteria has also highlighted prominent differences in the conformational flexibility among human and bacterial enzymes, yet unexplored for the design of selective bTS inhibitors.

These are promising results; however, more efforts are required to obtain drug candidates targeting bTS and FDTS enzymes.

Author Contributions: Writing—original draft preparation, C.P., L.L., G.T., and S.M.; writing—review and editing, C.P. and S.M.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in the manuscript:

dTMP	2′-deoxythymidine-5′-monophosphate
dUMP	2′-deoxyuridine-5′-monophosphate
TS	thymidylate synthase
DHFR	dihydrofolate reductase
FDTS	flavin dependent thymidylate synthase
CH_2H_4folate	N^5,N^{10}-methylene-5,6,7,8,-tetrahydrofolate
H_2folate	dihydrofolate
H_4folate	tetrahydrofolate
FAD	flavin adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
PDB	Protein Data Bank

References

1. Carreras, C.W.; Santi, D.V. The Catalytic Mechanism and Structure of Thymidylate Synthase. *Annu. Rev. Biochem.* **1995**, *64*, 721–762. [CrossRef]
2. Myllykallio, H.; Lipowski, G.; Leduc, D.; Filee, J.; Forterre, P.; Liebl, U. An Alternative Flavin-Dependent Mechanism for Thymidylate Synthesis. *Science* **2002**, *297*, 105–107. [CrossRef]
3. Kholodar, S.A.; Kohen, A. Noncovalent Intermediate of Thymidylate Synthase: Fact or Fiction? *J. Am. Chem. Soc.* **2016**, *138*, 8056–8059. [CrossRef] [PubMed]
4. Mishanina, T.V.; Yu, L.; Karunaratne, K.; Mondal, D.; Corcoran, J.M.; Choi, M.A.; Kohen, A. An unprecedented mechanism of nucleotide methylation in organisms containing thyX. *Science* **2016**, *351*, 507–510. [CrossRef]
5. Nazki, F.H.; Sameer, A.S.; Ganaie, B.A. Folate: Metabolism, genes, polymorphisms and the associated diseases. *Gene* **2014**, *533*, 11–20. [CrossRef]
6. Fivian-Hughes, A.S.; Houghton, J.; Davis, E.O. Mycobacterium tuberculosis thymidylate synthase gene thyX is essential and potentially bifunctional, while thyA deletion confers resistance to p-aminosalicylic acid. *Microbiology* **2012**, *158*, 308–318. [CrossRef] [PubMed]
7. Liu, X.-Q.; Yang, J. Bacterial Thymidylate Synthase with Intein, Group II Intron, and Distinctive ThyX Motifs. *J. Bacteriol.* **2004**, *186*, 6316–6319. [CrossRef]
8. Zhang, H.; Li, D.; Zhao, L.; Fleming, J.; Lin, N.; Wang, T.; Liu, Z.; Li, C.; Galwey, N.; Deng, J.; et al. Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance. *Nat. Genet.* **2013**, *45*, 1255–1260. [CrossRef] [PubMed]
9. Laxminarayan, R.; Duse, A.; Wattal, C.; Zaidi, A.K.M.; Wertheim, H.F.L.; Sumpradit, N.; Vlieghe, E.; Hara, G.L.; Gould, I.M.; Goossens, H.; et al. Antibiotic resistance—The need for global solutions. *Lancet Infect. Dis.* **2013**, *13*, 1057–1098. [CrossRef]
10. Sharma, V.K.; Johnson, N.; Cizmas, L.; McDonald, T.J.; Kim, H. A review of the influence of treatment strategies on antibiotic resistant bacteria and antibiotic resistance genes. *Chemosphere* **2016**, *150*, 702–714. [CrossRef]
11. Hiltunen, T.; Virta, M.; Laine, A.L. Antibiotic resistance in the wild: An eco-evolutionary perspective. *Philos. Trans. R. Soc. B Biol. Sci.* **2017**, *372*, 20160039. [CrossRef] [PubMed]
12. Martinez, J.L. General principles of antibiotic resistance in bacteria. *Drug Discov. Today: Technol.* **2014**, *11*, 33–39. [CrossRef] [PubMed]
13. Chernyshev, A.; Fleischmann, T.; Kohen, A. Thymidyl biosynthesis enzymes as antibiotic targets. *Appl. Microbiol. Biotechnol.* **2007**, *74*, 282–289. [CrossRef]
14. Choi, M.; Karunaratne, K.; Kohen, A. Flavin-Dependent Thymidylate Synthase as a New Antibiotic Target. *Molecules* **2016**, *21*, 654. [CrossRef] [PubMed]
15. Costi, M.P.; Ferrari, S.; Venturelli, A.; Calò, S.; Tondi, D.; Barlocco, D. Thymidylate synthase structure, function and implication in drug discovery. *Curr. Med. Chem.* **2005**, *12*, 2241–2258. [CrossRef] [PubMed]
16. Pozzi, C.; Ferrari, S.; Cortesi, D.; Luciani, R.; Stroud, R.M.; Catalano, A.; Costi, M.P.; Mangani, S. The structure of Enterococcus faecalis thymidylate synthase provides clues about folate bacterial metabolism. *Acta Cryst. D* **2012**, *68*, 1232–1241. [CrossRef]

17. Erlanson, D.A.; Braisted, A.C.; Raphael, D.R.; Randal, M.; Stroud, R.M.; Gordon, E.M.; Wells, J.A. Site-directed ligand discovery. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 9367–9372. [CrossRef]
18. Sapienza, P.J.; Falk, B.T.; Lee, A.L. Bacterial Thymidylate Synthase Binds Two Molecules of Substrate and Cofactor without Cooperativity. *J. Am. Chem. Soc.* **2015**, *137*, 14260–14263. [CrossRef]
19. Finer-Moore, J.S.; Lee, T.T.; Stroud, R.M. A Single Mutation Traps a Half-Sites Reactive Enzyme in Midstream, Explaining Asymmetry in Hydride Transfer. *Biochemistry* **2018**, *57*, 2786–2795. [CrossRef]
20. Salo-Ahen, O.M.H.; Tochowicz, A.; Pozzi, C.; Cardinale, D.; Ferrari, S.; Boum, Y.; Mangani, S.; Stroud, R.M.; Saxena, P.; Myllykallio, H.; et al. Hotspots in an Obligate Homodimeric Anticancer Target. Structural and Functional Effects of Interfacial Mutations in Human Thymidylate Synthase. *J. Med. Chem.* **2015**, *58*, 3572–3581. [CrossRef]
21. Islam, Z.; Gurevic, I.; Strutzenberg, T.S.; Ghosh, A.K.; Iqbal, T.; Kohen, A. Bacterial versus human thymidylate synthase: Kinetics and functionality. *PLoS ONE* **2018**, *13*, e0196506. [CrossRef] [PubMed]
22. Kawase, S.; Cho, S.-W.; Rozelle, J.; Stroud, R.M.; Finer-Moore, J.; Santi, D.V. Replacement set mutagenesis of the four phosphate-binding arginine residues of thymidylate synthase. *Protein Eng. Des. Sel.* **2000**, *13*, 557–563. [CrossRef]
23. Sotelo-Mundo, R.R.; Ciesla, J.; Dzik, J.M.; Rode, W.; Maley, F.; Maley, G.F.; Hardy, L.W.; Montfort, W.R. Crystal Structures of Rat Thymidylate Synthase Inhibited by Tomudex, a Potent Anticancer Drug. *Biochemistry* **1999**, *38*, 1087–1094. [CrossRef]
24. Strop, P.; Changchien, L.; Maley, F.; Montfort, W.R. Crystal structures of a marginally active thymidylate synthase mutant, Arg 126–>Glu. *Protein Sci.* **1997**, *6*, 2504–2511. [CrossRef]
25. Sotelo-Mundo, R.R.; Changchien, L.; Maley, F.; Montfort, W.R. Crystal structures of thymidylate synthase mutant R166Q: Structural basis for the nearly complete loss of catalytic activity. *J. Biochem. Mol. Toxicol.* **2006**, *20*, 88–92. [CrossRef]
26. Lopez-Zavala, A.A.; Guevara-Hernandez, E.; Vazquez-Lujan, L.H.; Sanchez-Paz, A.; Garcia-Orozco, K.D.; Contreras-Vergara, C.A.; Lopez-Leal, G.; Arvizu-Flores, A.A.; Ochoa-Leyva, A.; Sotelo-Mundo, R.R. A novel thymidylate synthase from the Vibrionales, Alteromonadales, Aeromonadales, and Pasteurellales (VAAP) clade with altered nucleotide and folate binding sites. *PeerJ* **2018**, *6*, e5023. [CrossRef]
27. Reyes, C.L.; Sage, C.R.; Rutenber, E.E.; Nissen, R.M.; Finer-Moore, J.S.; Stroud, R.M. Inactivity of N229A thymidylate synthase due to water-mediated effects: Isolating a late stage in methyl transfer11Edited by D. Rees. *J. Mol. Biol.* **1998**, *284*, 699–712. [CrossRef]
28. Wang, Z.; Abeysinghe, T.; Finer-Moore, J.S.; Stroud, R.M.; Kohen, A. A Remote Mutation Affects the Hydride Transfer by Disrupting Concerted Protein Motions in Thymidylate Synthase. *J. Am. Chem. Soc.* **2012**, *134*, 17722–17730. [CrossRef]
29. Arvizu-Flores, A.A.; Sugich-Miranda, R.; Arreola, R.; Garcia-Orozco, K.D.; Velazquez-Contreras, E.F.; Montfort, W.R.; Maley, F.; Sotelo-Mundo, R.R. Role of an invariant lysine residue in folate binding on Escherichia coli thymidylate synthase: Calorimetric and crystallographic analysis of the K48Q mutant. *Int. J. Biochem. Cell Biol.* **2008**, *40*, 2206–2217. [CrossRef]
30. Spencer, H.T.; Villafranca, J.E.; Appleman, J.R. Kinetic Scheme for Thymidylate Synthase from Escherichia coli: Determination from Measurements of Ligand Binding, Primary and Secondary Isotope Effects, and Pre-Steady-State Catalysis. *Biochemistry* **1997**, *36*, 4212–4222. [CrossRef]
31. Stroud, R.M.; Finer-Moore, J.S. Conformational dynamics along an enzymatic reaction pathway: Thymidylate synthase, "the movie". *Biochemistry* **2003**, *42*, 239–247. [CrossRef]
32. Kanaan, N.; Ferrer, S.; Martí, S.; Garcia-Viloca, M.; Kohen, A.; Moliner, V. Temperature Dependence of the Kinetic Isotope Effects in Thymidylate Synthase. A Theoretical Study. *J. Am. Chem. Soc.* **2011**, *133*, 6692–6702. [CrossRef] [PubMed]
33. Wang, Z.; Ferrer, S.; Moliner, V.; Kohen, A. QM/MM Calculations Suggest a Novel Intermediate Following the Proton Abstraction Catalyzed by Thymidylate Synthase. *Biochemistry* **2013**, *52*, 2348–2358. [CrossRef]
34. Kaiyawet, N.; Lonsdale, R.; Rungrotmongkol, T.; Mulholland, A.J.; Hannongbua, S. High-Level QM/MM Calculations Support the Concerted Mechanism for Michael Addition and Covalent Complex Formation in Thymidylate Synthase. *J. Chem. Theory Comput.* **2015**, *11*, 713–722. [CrossRef] [PubMed]
35. Islam, Z.; Strutzenberg, T.S.; Gurevic, I.; Kohen, A. Concerted versus Stepwise Mechanism in Thymidylate Synthase. *J. Am. Chem. Soc.* **2014**, *136*, 9850–9853. [CrossRef] [PubMed]
36. Islam, Z.; Strutzenberg, T.S.; Ghosh, A.K.; Kohen, A. Activation of Two Sequential H Transfers in the Thymidylate Synthase Catalyzed Reaction. *ACS Catal.* **2015**, *5*, 6061–6068. [CrossRef]

37. Maley, F.; Pedersen-Lane, J.; Changchien, L. Complete Restoration of Activity to Inactive Mutants of Escherichia coli Thymidylate Synthase: Evidence that E. coli Thymidylate Synthase is a Half-the-Sites Activity Enzyme. *Biochemistry* **1995**, *34*, 1469–1474. [CrossRef]
38. Chen, C.-H.; Davis, R.A.; Maley, F. Thermodynamic Stabilization of Nucleotide Binding to Thymidylate Synthase by A Potent Benzoquinazoline Folate Analogue Inhibitor. *Biochemistry* **1996**, *35*, 8786–8793. [CrossRef]
39. Sapienza, P.J.; Lee, A.L. Widespread Perturbation of Function, Structure, and Dynamics by a Conservative Single-Atom Substitution in Thymidylate Synthase. *Biochemistry* **2016**, *55*, 5702–5713. [CrossRef]
40. Falk, B.T.; Sapienza, P.J.; Lee, A.L. Chemical shift imprint of intersubunit communication in a symmetric homodimer. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9533–9538. [CrossRef]
41. Pozzi, C.; Ferrari, S.; Luciani, R.; Tassone, G.; Costi, M.P.; Mangani, S. Structural Comparison of Enterococcus faecalis and Human Thymidylate Synthase Complexes with the Substrate dUMP and Its Analogue FdUMP Provides Hints about Enzyme Conformational Variabilities. *Molecules* **2019**, *24*, 1257. [CrossRef]
42. Sampathkumar, P.; Turley, S.; Ulmer, J.E.; Rhie, H.G.; Sibley, C.H.; Hol, W.G.J. Structure of the Mycobacterium tuberculosis Flavin Dependent Thymidylate Synthase (MtbThyX) at 2.0Å Resolution. *J. Mol. Biol.* **2005**, *352*, 1091–1104. [CrossRef]
43. Wang, K.; Wang, Q.; Chen, J.; Chen, L.; Jiang, H.; Shen, X. Crystal structure and enzymatic characterization of thymidylate synthase X from Helicobacter pylori strain SS1. *Protein Sci.* **2011**, *20*, 1398–1410. [CrossRef] [PubMed]
44. Baugh, L.; Phan, I.; Begley, D.W.; Clifton, M.C.; Armour, B.; Dranow, D.M.; Taylor, B.M.; Muruthi, M.M.; Abendroth, J.; Fairman, J.W.; et al. Increasing the structural coverage of tuberculosis drug targets. *Tuberculosis* **2015**, *95*, 142–148. [CrossRef]
45. Sampathkumar, P.; Turley, S.; Sibley, C.H.; Hol, W.G.J. NADP+ Expels both the Co-factor and a Substrate Analog from the Mycobacterium tuberculosis ThyX Active Site: Opportunities for Anti-bacterial Drug Design. *J. Mol. Biol.* **2006**, *360*, 1–6. [CrossRef] [PubMed]
46. Koehn, E.M.; Perissinotti, L.L.; Moghram, S.; Prabhakar, A.; Lesley, S.A.; Mathews, I.I.; Kohen, A. Folate binding site of flavin-dependent thymidylate synthase. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 15722–15727. [CrossRef] [PubMed]
47. Ferrari, S.; Severi, L.; Pozzi, C.; Quotadamo, A.; Ponterini, G.; Losi, L.; Marverti, G.; Costi, M.P. Chapter Seventeen—Human Thymidylate Synthase Inhibitors Halting Ovarian Cancer Growth. In *Vitamins and Hormones*; Litwack, G., Ed.; Academic Press: Cambridge, MA, USA, 2018; Volume 107, pp. 473–513.
48. Chen, D.; Jansson, A.; Sim, D.; Larsson, A.; Nordlund, P. Structural analyses of human thymidylate synthase reveal a site that may control conformational switching between active and inactive states. *J. Biol. Chem.* **2017**, *292*, 13449–13458. [CrossRef]
49. Mangani, S.; Cancian, L.; Leone, R.; Pozzi, C.; Lazzari, S.; Luciani, R.; Ferrari, S.; Costi, M.P. Identification of the Binding Modes of N-Phenylphthalimides Inhibiting Bacterial Thymidylate Synthase through X-Ray Crystallography Screening. *J. Med. Chem.* **2011**, *54*, 5454–5467. [CrossRef]
50. Catalano, A.; Luciani, R.; Carocci, A.; Cortesi, D.; Pozzi, C.; Borsari, C.; Ferrari, S.; Mangani, S. X-ray crystal structures of Enterococcus faecalis thymidylate synthase with folate binding site inhibitors. *Eur. J. Med. Chem.* **2016**, *123*, 649–664. [CrossRef]
51. Ferrari, S.; Ingrami, M.; Soragni, F.; Wade, R.C.; Costi, M.P. Ligand-based discovery of N-(1,3-dioxo-1H,3H-benzo[de]isochromen-5-yl)-carboxamide and sulfonamide derivatives as thymidylate synthase A inhibitors. *Bioorgan. Med. Chem. Lett.* **2013**, *23*, 663–668. [CrossRef]
52. Ferrari, S.; Calò, S.; Leone, R.; Luciani, R.; Costantino, L.; Sammak, S.; Di Pisa, F.; Pozzi, C.; Mangani, S.; Costi, M.P. 2′-Deoxyuridine 5′-Monophosphate Substrate Displacement in Thymidylate Synthase through 6-Hydroxy-2H-naphtho[1,8-bc]furan-2-one Derivatives. *J. Med. Chem.* **2013**, *56*, 9356–9360. [CrossRef] [PubMed]
53. Fargualy, A.M.; Habib, N.S.; Ismail, K.A.; Hassan, A.M.M.; Sarg, M.T.M. Synthesis, biological evaluation and molecular docking studies of some pyrimidine derivatives. *Eur. J. Med. Chem.* **2013**, *66*, 276–295. [CrossRef]
54. Palepu, N.R.; Nongbri, S.L.; Premkumar, J.R.; Verma, A.K.; Bhattacharjee, K.; Joshi, S.R.; Forbes, S.; Mozharivskyj, Y.; Thounaojam, R.; Aguan, K.; et al. Synthesis and evaluation of new

salicylaldehyde-2-picolinylhydrazone Schiff base compounds of Ru(II), Rh(III) and Ir(III) as in vitro antitumor, antibacterial and fluorescence imaging agents. *J. Biol. Inorg. Chem.* **2015**, *20*, 619–638. [CrossRef] [PubMed]

55. Sayre, P.H.; Finer-Moore, J.S.; Fritz, T.A.; Biermann, D.; Gates, S.B.; MacKellar, W.C.; Patel, V.F.; Stroud, R.M. Multi-targeted antifolates aimed at avoiding drug resistance form covalent closed inhibitory complexes with human and Escherichia coli thymidylate synthases. *J. Mol. Biol.* **2001**, *313*, 813–829. [CrossRef]
56. Basta, T.; Boum, Y.; Briffotaux, J.; Becker, H.F.; Lamarre-Jouenne, I.; Lambry, J.-C.; Skouloubris, S.; Liebl, U.; Graille, M.; van Tilbeurgh, H.; et al. Mechanistic and structural basis for inhibition of thymidylate synthase ThyX. *Open Biol.* **2012**, *2*, 120120. [CrossRef]
57. Skouloubris, S.; Djaout, K.; Lamarre, I.; Lambry, J.-C.; Anger, K.; Briffotaux, J.; Liebl, U.; Reuse, H.; Myllykallio, H. Targeting of Helicobacter pylori thymidylate synthase ThyX by non-mitotoxic hydroxy-naphthoquinones. *Open Biol.* **2015**, *5*, 150015. [CrossRef]
58. Luciani, R.; Saxena, P.; Surade, S.; Santucci, M.; Venturelli, A.; Borsari, C.; Marverti, G.; Ponterini, G.; Ferrari, S.; Blundell, T.L.; et al. Virtual Screening and X-ray Crystallography Identify Non-Substrate Analog Inhibitors of Flavin-Dependent Thymidylate Synthase. *J. Med. Chem.* **2016**, *59*, 9269–9275. [CrossRef]
59. Abu El Asrar, R.; Margamuljana, L.; Klaassen, H.; Nijs, M.; Marchand, A.; Chaltin, P.; Myllykallio, H.; Becker, H.F.; De Jonghe, S.; Herdewijn, P.; et al. Discovery of a new Mycobacterium tuberculosis thymidylate synthase X inhibitor with a unique inhibition profile. *Biochem. Pharmacol.* **2017**, *135*, 69–78. [CrossRef] [PubMed]
60. Hunter, J.H.; Gujjar, R.; Pang, C.K.T.; Rathod, P.K. Kinetics and Ligand-Binding Preferences of Mycobacterium tuberculosis Thymidylate Synthases, ThyA and ThyX. *PLoS ONE* **2008**, *3*, e2237. [CrossRef] [PubMed]
61. Singh, V.; Brecik, M.; Mukherjee, R.; Evans, J.C.; Svetlíková, Z.; Blaško, J.; Surade, S.; Blackburn, J.; Warner, D.F.; Mikušová, K.; et al. The Complex Mechanism of Antimycobacterial Action of 5-Fluorouracil. *Chem. Biol.* **2015**, *22*, 63–75. [CrossRef]
62. Kögler, M.; Vanderhoydonck, B.; De Jonghe, S.; Rozenski, J.; Van Belle, K.; Herman, J.; Louat, T.; Parchina, A.; Sibley, C.; Lescrinier, E.; et al. Synthesis and Evaluation of 5-Substituted 2′-deoxyuridine Monophosphate Analogues As Inhibitors of Flavin-Dependent Thymidylate Synthase in Mycobacterium tuberculosis. *J. Med. Chem.* **2011**, *54*, 4847–4862. [CrossRef]
63. Kögler, M.; Busson, R.; De Jonghe, S.; Rozenski, J.; Van Belle, K.; Louat, T.; Munier-Lehmann, H.; Herdewijn, P. Synthesis and Evaluation of 6-Aza-2′-deoxyuridine Monophosphate Analogs as Inhibitors of Thymidylate Synthases, and as Substrates or Inhibitors of Thymidine Monophosphate Kinase in Mycobacterium tuberculosis. *Chem. Biodivers.* **2012**, *9*, 536–556. [CrossRef] [PubMed]
64. Parchina, A.; Froeyen, M.; Margamuljana, L.; Rozenski, J.; De Jonghe, S.; Briers, Y.; Lavigne, R.; Herdewijn, P.; Lescrinier, E. Discovery of an Acyclic Nucleoside Phosphonate that Inhibits Mycobacterium tuberculosis ThyX Based on the Binding Mode of a 5-Alkynyl Substrate Analogue. *ChemMedChem* **2013**, *8*, 1373–1383. [CrossRef] [PubMed]
65. Alexandrova, L.A.; Chekhov, V.O.; Shmalenyuk, E.R.; Kochetkov, S.N.; El-Asrar, R.A.; Herdewijn, P. Synthesis and evaluation of C-5 modified 2′-deoxyuridine monophosphates as inhibitors of M. tuberculosis thymidylate synthase. *Bioorgan. Med. Chem.* **2015**, *23*, 7131–7137. [CrossRef] [PubMed]
66. Shmalenyuk, E.R.; Chernousova, L.N.; Karpenko, I.L.; Kochetkov, S.N.; Smirnova, T.G.; Andreevskaya, S.N.; Chizhov, A.O.; Efremenkova, O.V.; Alexandrova, L.A. Inhibition of Mycobacterium tuberculosis strains H37Rv and MDR MS-115 by a new set of C5 modified pyrimidine nucleosides. *Bioorgan. Med. Chem.* **2013**, *21*, 4874–4884. [CrossRef]
67. Negrya, S.D.; Efremenkova, O.V.; Solyev, P.N.; Chekhov, V.O.; Ivanov, M.A.; Sumarukova, I.G.; Karpenko, I.L.; Kochetkov, S.N.; Alexandrova, L.A. Novel 5-substituted derivatives of 2′-deoxy-6-azauridine with antibacterial activity. *J. Antibiot.* **2019**, 1. [CrossRef]

Part III

Structural insight into hYAP1-hTEAD4 protein-protein interaction: a novel target for cancer treatment

Chapter 7

Introduction

The Hippo pathway is an evolutionarily conserved intracellular signaling cascade having a central role in organ size control and tissue homeostasis, leading both development and regeneration of different organisms [44, 139, 140]. The Hippo pathway has a critical function in regulating fundamental biological processes, as cell growth, differentiation, survival and apoptosis [44, 139, 141, 142]. The components of the pathway were originally identified in *Drosophila*; then, orthologues with similar roles in the signaling network were discovered in mammals [139, 143]. The core of the Hippo pathway consists of several serine/threonine kinases activated through a phosphorylation cascade (*Figure 7.1*). Once the pathway is turned on, mammalian Ste20-like 1 and 2 (MST1/2) kinases are phosphorylated and, after this modification, they phosphorylate the Large Tumor Suppressor Kinases (LATS1/2), which interact with the adaptor protein Mps1-binder-related 1 (MOB1) (*Figure 7.1*). At this stage, the LATS 1/2 – MOB1 complex phosphorylates the Yes associated protein (YAP), resulting in its cytoplasmic retention and subsequent degradation by the ubiquitin-proteasome system [44, 139, 141, 144] (*Figure 7.1* and *Figure 7.2*). Conversely, when the Hippo pathway is switched off, YAP is not phosphorylated by LATS1/2 kinases, ant can translocate into the nucleus where it interacts with different partners, mainly the transcriptional enhanced associate domain (TEAD 1-4) transcription factors (*Figure 7.1* and *Figure 7.2*) [44, 139, 141, 145].

Thus, TEADs are the distal mediators of the Hippo signaling and their interaction with YAP is essential for the expression of the Hippo pathway-related genes (e.g. *Myc*, *Axl*, *EGFR*, *PD-L1* and *EGFR*) (*Figure 7.2*) [44, 146, 147]. Consequently, YAP and TEADs are main downstream effectors of this signaling pathway. In human cells, the expression of pro-proliferation, pro-survival and anti-apoptosis target genes is triggered upon hYAP:hTEADs complex formation [139, 143, 148, 149].

hTEAD proteins are a family of transcription factors which control the expression of a wide set of genes related to proliferation, differentiation, and apoptosis [150]. In humans, the family consists of four different genes, which encode for the four homologous members of the family, named hTEAD1-4

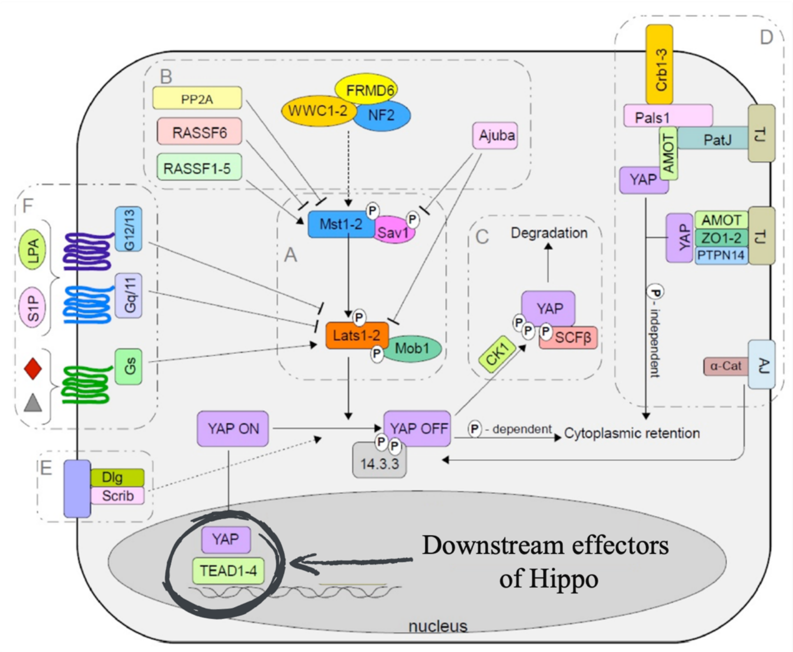


Figure 7.1: The Hippo signaling pathway in mammals, which consists of a serine/threonine phosphorylation cascade, leading to the translocation of YAP into the nucleus where it interacts with TEADs. The YAP:TEAD1-4 complex, acting as downstream effector of the mammalian Hippo pathway, is underlined in the Figure [44].

(*Figure 7.3*). The isoforms show tissue and development-specific expression patterns, and they can be expressed simultaneously in the same tissue [148]. In this work, we focused on the hTEAD4 isoform which is the most abundant in human cells and plays a role in a plethora of cancers, as outlined below [151, 152, 153]. hTEAD4 is a nuclear protein consisting of a single polypeptide chain with 434 amino acid residues and a molecular weight of approximatively 49 kDa [148]. hTEAD4 shares a high homology sequence (76.5%, 64.27% and 71.06% sequence homology respect to hTEAD1,hTEAD2,hTEAD3 and hTEAD4 respectively; *Figure 7.3*) and a common overall architecture with the other family members [146, 148, 150].

The N-terminus of hTEAD4 is constituted by a TEA DNA-binding domain (DNA-BD), required for its binding to DNA cis elements; while the YAP/TAZ binding domain (generally referred to as YBD) has been mapped at the hTEAD C-terminus [148, 150, 154] (*Figure 7.4*). Here, we focus on the hTEAD4 YBD, a

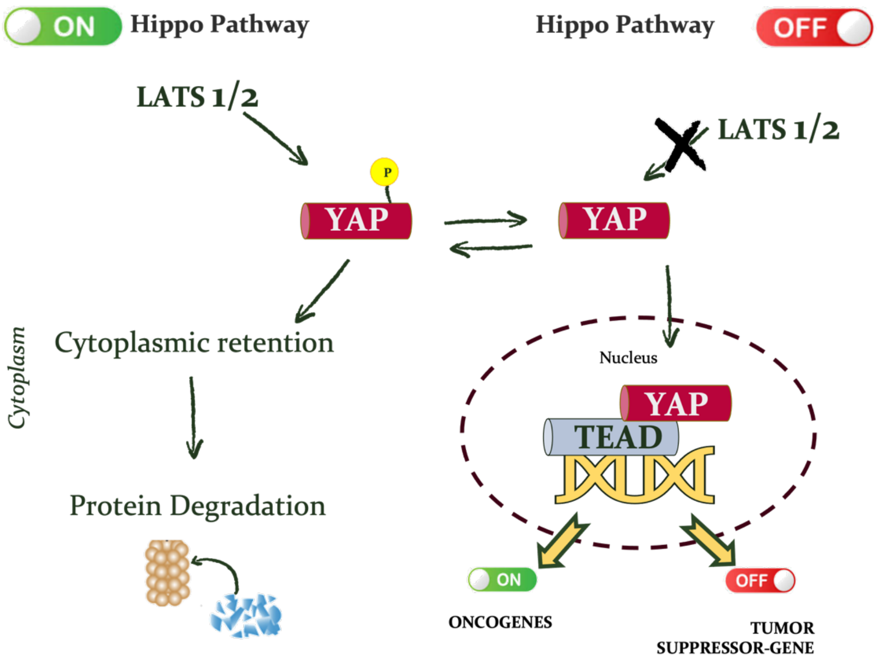


Figure 7.2: Schematic representation of Hippo signaling events in both ON and OFF pathway states, which determine the nucleus/cytoplasm shuttle of YAP.

```
hTEAD4     ---MEGTAGTITSNEWSSPTSPEGSTASGGSQALDKPIDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEGKMYGRNELIARYIKLRTGKTRTRKQVSSHIQVLARRKAREIQAKL (117)
hTEAD1     ----------MEPSSWSGSESPAENM-ERMSDSADKPIDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEGKMYGRNELIARYIKLRTGKTRTRKQVSSHIQVLARRKSRDFHSKL (109)
hTEAD2     MGEPRAGAALDDGSGWTGSEEGSE-EGTGGSEGAGGDGGPDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEGKMYGRNELIARYIKLRTGKTRTRKQVSSHIQVLARRKSREIQSKL (119)
hTEAD3     ----------MASNSWNASSSPGEAR-EDGPEGLDKGLDNDAEGVWSPDIEQSFQEALAIYPPCGRRKIILSDEGKMYGRNELIARYIKLRTGKTRTRKQVSSHIQVLARKKVREYQVGI (109)
                        . *..  .          :. .   . **********************************************************************:* *: :  :

hTEAD4     K----DQAAKDKALQSMAAMSSAQIISATAFHSSMALAR---GPGRPAVSGFWQGAL—PGQAGTSHDVKPFSQQTYAVQP-PL--PLPGFESPAGPAPS--PSAPPAPPWQGRSVASSKL (224)
hTEAD1     K----DQTAKDKALQHMAAMSSAQIVSATAIHNKLGLPG-IPRPTFPGAPGFWPGMI-QTGQPGSSQDVKPFVQQAYPIQP-AVTAPIPGFEPASAP------APSVPAWQGRSIGTTKL (216)
hTEAD2     K----DQVSKDKAFQTMATMSSAQLISAPSLQAKLGPTGPQAS----ELFQFWSGG---SGPPWNVPDVKPFSQTPFTLSLTPPSTDLPGYEPPQALSPLPPPTPSPPAWQARGLGTARL (228)
hTEAD3     KAMNLDQVSKDKALQSMASMSSAQIVSASVLQNKFSPPSPLPQAVFSTSSRFWSSPPLLGQQPGPSQDIKPFAQPAYPIQP-PLPPTLSSYEPLA-PLPS--AAASVPVWQDRTIASSRL (225)
           *    **.:****:* **:*****::**  :: .:.                **  .               *:*** *  : :.        : .:*               :   * ** * :.::*

hTEAD4     WMLEFSAFLEQQQDPDTYNKHLFVHIGQSSPSYSDPYLEAVDIRQIYDKFPEKKGGLKDLFERGPSNAFFLVKFWADLNTNIED---------EGSSFYGVSSQYESPENMIITCSTKVC (335)
hTEAD1     RLVEFSAFLEQQRDPDSYNKHLFVHIGHANHSYSDPLLESVDIRQIYDKFPEKKGGLKELFGKGPQNAFFLVKFWADLNCNIQD---------DAGAFYGVTSQYESSENMTVTCSTKVC (327)
hTEAD2     QLVEFSAFVEPPDAVDSYQRHLFVHISQHCPSPGAPPLESVDVRQIYDKFPEKKGGLRELYDRGPPHAFFLVKFWADLNWGPSGEEAGAGGSISSGGFYGVSSQYESLEHMTLTCSSKVC (348)
hTEAD3     RLLEYSAFMEVQRDPDTYSKHLFVHIGQTNPAFSDPPLEAVDVRQIYDKFPEKKGGLKELYEKGPPNAFFLVKFWADLNSTIQE---------GPGAFYGVSSQYSSADSMTISVSTKVC (336)
            ::*:***:*     *:*.:******.:   : . * **:**:**************::*: :** :************   .              ..****:***.* : * :: *:***

hTEAD4     SFGKQVVEKVETEYARYENGHYSYRIHRSPLCEYMINFIHKLKHLPEKYMMNSVLENFTILQVVTNRDTQETLLCIAYVFEVSASEHGAQHHIYRLVKE (434)
hTEAD1     SFGKQVVEKVETEYARFENGRFVYRINRSPMCEYMINFIHKLKHLPEKYMMNSVLENFTILLVVTNRDTQETLLCMACVFEVSNSEHGAQHHIYRLVKD (426)
hTEAD2     SFGKQVVEKVETERAQLEDGRFVYRLLRSPMCEYLVNFLHKLRQLPERYMMNSVLENFTILQVVTNRDTQELLLCTAYVFEVSTSERGAQHHIYRLVRD (447)
hTEAD3     SFGKQVVEKVETEYARLENGRFVYRIHRSPMCEYMINFIHKLKHLPEKYMMNSVLENFTILQVVTSRDSQETLLVIAFVFEVSTSEHGAQHHVYKLVKD (435)
           ************* *: *:*:: **: ***:***::**:***::***:************ ***.**:** **  * ***** **:*****:*:**::
```

Figure 7.3: Sequence alignment among the four hTEAD isoforms (UniProtKB codes, hTEAD4: Q15561; hTEAD1: P28347; hTEAD2: Q15562; hTEAD3: Q99594) performed with Clustal Omega. The asterisk (*) indicates positions which have a single, fully conserved residue; the symbol colon (:) indicates conservation between groups of strongly similar properties; the period (.) indicates conservation between groups of weakly similar properties.

transactivation domain mediating the interaction with its protein coactivator hYAP1. Although being transcriptional factors, hTEADs are not able to induce gene transcription by themselves, but coactivator interactions are required to modulate gene expression [155]. Furthermore, all hTEADs have a deep hydrophobic central pocket in their YBD, where palmitic or myristic acid can bind under physiological conditions (*Figure 7.4*), establishing a covalent linkage with a conserved cysteine residue (Cys359 in hTEAD1, Cys380 in hTEAD2, Cys371 in hTEAD3 and Cys367 in hTEAD4) [156, 157, 158]. It has been demonstrated that hTEADs can undergo spontaneous acylation, in presence of physiological concentrations of either palmitoyl-CoA or myristoyl-CoA [156, 157]. The role of hTEAD4 acylation has not been fully clarified yet. Firstly, it was suggested that hTEAD4 acylation is important for the interaction with hYAP [157]. However, recent biochemical studies have shown that the hYAP1:hTEAD4 interaction is not affected by the acylation of the latter partner, which seems stabilized upon this modification [156].

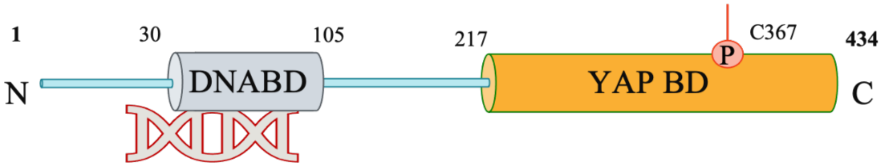


Figure 7.4: Domain architecture of hTEAD4. TEADs consist of a TEA DNA binding domain (DNA BD, in grey) and a YAP/TAZ binding domain (YAP BD, in orange). Cys367 (or C367), which is the site of hTEAD4 acylation, is displayed in the Figure. Residues included in each domain are shown.

On the other hand, hYAP is a transcriptional co-activator without an intrinsic DNA-BD that contains a key activation domain for the association with the cellular transcriptional machinery [145, 159]. hYAP shuttles between the cytoplasm and the nucleus where it interacts with hTEADs and drives the expression of Hippo pathway-target genes. The function of hYAP depends on its ability to recruit the cellular transcriptional machinery following the binding to hTEAD transcriptional factors which are bound to the gene promoters. It is a modular protein having 504 amino acids and consists of several structural domains (*Figure 7.5*). As mapped by *Vassilev and coworkers*, the N-terminus of hYAP is a proline-rich region, followed by the TEAD-binding domain (TBD, including amino acids 50-171) [154], two WW small modular domains (named WW1 and WW2, including residues 171-204 and 230-263, respectively) [160], a

SH3-binding motif, a coiled-coil (CC) region, and a transcription activation domain (TAD, determinant for its transcriptional co-regulatory role) (*Figure 7.5*) [154, 160, 161, 162]. hYAP is an intrinsically disordered protein which seems to structure upon the interaction with hTEADs [163], however, to date, the structural information on hYAP is limited to the TBD peptide 50-99 in heterodimeric complexes with hTEADs [156, 161, 163].

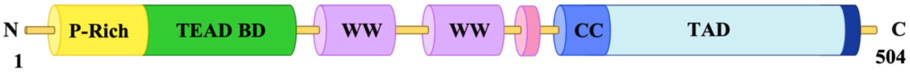


Figure 7.5: Schematic representation of domain organization of the modular protein hYAP. The TBD, WW1, WW2, CC, TAD domains are shown.

The biological processes controlled by the Hippo pathway are responsible for multiple oncogenic hallmarks. Indeed, Hippo-pathway dysfunctions, leading to uncontrolled proliferation and resistance to programmed cell death, are well documented in several human malignancies [44, 139, 143, 149, 164]. Alterations of the Hippo signaling account also for cancer progression, by modulating cell migration, invasion, angiogenesis, immune-deficiency, and metastasis, and for drug resistance [139, 143, 164]. The main proteins responsible for Hippo pathway dysfunctions in cancer are hYAP and hTEADs, thus classified as oncoproteins [164]. The hYAP:hTEAD complex plays a dual role in cancer, promoting the expression of oncogenes and concomitantly blocking that of tumor suppressor genes (*Figure 7.2*). Increased hYAP protein levels and nuclear localization are reported in various human cancers, such as osteosarcoma (OS), mesothelioma, Renal Cell Carcinoma (RCC), gastric, endometrial, lung and liver carcinoma [141, 149, 151, 152, 153, 165]. Analogously, increased expression of hTEAD is associated with poor prognosis in gastric, breast, colorectal, RCC and prostate cancers [147, 150, 166, 167, 168]. Increasing attention has been recently focused on the hTEAD4 isoform which has been associated with high pathogenicity of head and neck squamous cell carcinoma (HNSCC) [169], and with poor prognosis in lung adenocarcinoma and breast cancer [170]. Measuring the expression levels of hYAP and hTEADs can thus represent a valid prognostic marker in the tumor subsets where their overexpression has been correlated with poor therapeutic outcomes[171]. Furthermore, the re-establishment of physiological levels of hYAP and hTEAD proteins has been directly correlated with cellular sensitization to chemotherapy and radiotherapy [172]. In this scenario, targeting the hYAP:hTEADs complex represents an emerging, promising therapeutic strategy for cancer treatment [44, 149]. Very few hYAP:hTEADs inhibitors have been reported so far and the development of new molecules targeting this protein complex remains a challenge [44, 164, 173], due also the limited structural information available on this PPI [44]. Through a drug repurposing approach, Verteporfin (VP) has been recently proposed as potential modulator of the Hippo pathway terminal effectors [173, 174, 175]. Two main approaches have been proposed to disrupt the hYAP:hTEAD complex: one relies on the development of inhibitors directly binding at the hYAP-hTEAD4 interface, perturbing the complex formation, whereas the second is based on allosteric inhibitors targeting the hTEAD acylation pocket. At variance with the former approach whose efficacy is still under investigation [44, 176, 177], the second seems not effective. Indeed, molecules targeting the the hTEAD acylation pocket, like flufenamic and niflumic acids, does not alter the hYAP:hTEAD PPI [178]. As discussed in part I, achieving an efficient PPI disruption by small molecules is very difficult due to the large interacting surfaces connecting the two partner proteins [26, 28]. On the other hand, peptides seem interesting candidates for this purpose; indeed, starting from the hYAP Ω-loop amino acid sequence, *Zhang and colleagues* developed a potent cyclic peptide acting as hYAP mimetic [179]. The searching for novel hYAP:hTEAD PPI inhibitors and innovative inhibition strategies is thus still a challenge. Currently, the structural information available on hYAP and hTEAD4 is very limited and the structures of the full-length proteins have not been determined yet. More specifically, the information on this PPI is restricted to the partner domains, in the structure of the heterodimeric complex between the hYAP1 TBD (peptide 50-99) and the hTEAD4 YBD. The X-ray crystallographic studies performed so far on the hYAP:hTEADs complexes has provided basis to explain the interaction of these partner protein. In *2010*, the structure of the hYAP:hTEAD1 complex (PDB id 3KYS [161]) was firstly described, allowing the scientific community to gain valuable information on the main features of this PPI, relying on the presence of three interacting areas [161] (*Figure 7.6 A*). Next, the structural characterization of the hYAP:hTEAD4 complex (PDB ids 5OAQ [156] and 6GE3 [163]) showed some differences, displaying only two of these interface regions (*Figure 7.6 B*).

The purpose of this study is to expand the current structural and functional knowledge on the hYAP:hTEAD4 interaction and to explore the mechanism regulating this PPI. This information is essential for the development of innovative anticancer treatments. On this purpose we have developed reliable, effective protocols for expression and purification of hTEAD4 YBD and for the co-expression and co-purification of hTEAD4 YBD in complex with three hYAP TBD fragments (named S, M, and L complexes). Notably, the structure of complex S has been solved by means of X-ray crystallography, revealing new targetable interfaces between the partner proteins. Moreover, we have also co-expressed and co-purified the full-length hYAP in complex with hTEAD4 YBD and performed preliminary structural characterizations to obtain deeper insights on this key PPI within the Hippo signaling pathway. In this work, we report the studies on the isolated hTEAD4 YBD and four protein-protein complexes in which hTEAD4 YBD interacts either with the three hYAP TBD fragments 50-120 (S), 50-140 (M), and 50-171 (L), or with the full length hYAP. For the investigation of the target proteins, different genetic constructs for recombinant protein expression were generated using sub-cloning and site-direct mutagenesis techniques. Single proteins and protein-protein complexes (*Table 7.1* reports the plasmids, inserts, recombinant vectors and proteins or protein-protein complexes), were expressed, purified, and characterized in this thesis work.

Plasmid	Restriction sites	Insert name	Recombinant vector	Recombinant protein or protein-proten complex
pGEX-4T1	BamHI/XhoI	hTEAD4 YBD	pGEX-4T1-GST-hTEAD4 YBD	hTEAD4 YBD (217-434)
pET15b	NdeI/BamHI	hTEAD4 YBD	pET15b-His^6 -hTEAD4 YBD	hTEAD4 YBD (217-434)
pQE80-L	BamHI/HindIII	hYAP1	pQE80L-His^6-hYAP1	hYAP1 full lenght(1-504)
pRSFDuet-1	BamHI/HindIII NdeI/XhoI	hYAP1 TBD S hTEAD4 YBD	pRSFDuet-1-His^6-hYAP1 TBD S: hTEAD4 YBD	hYAP1 TBD (50-120): hTEAD4 YBD (217-434) S complex
pRSFDuet-1	BamHI/HindIII NdeI/XhoI	hYAP1 TBD M hTEAD4 YBD	pRSFDuet-1-His^6-hYAP1 TBD M: hTEAD4 YBD	hYAP1 TBD (50-140): hTEAD4 YBD (217-434) M complex
pRSFDuet-1	BamHI/HindIII NdeI/XhoI	hYAP1 TBD L hTEAD4 YBD	pRSFDuet-1-His^6- hYAP1 TBD L: hTEAD4 YBD	hYAP1 TBD (50-171): hTEAD4 YBD (217-434) L complex
pRSFDuet-1	BamHI/HindIII NdeI/XhoI	hYAP1 hTEAD4 YBD	pRSFDuet-1- His^6-hYAP1 : hTEAD4 YBD	hYAP1 TBD (1-504): hTEAD4 YBD (217-434) XL complex

Table 7.1: Summary of the plasmids, inserts, recombinant vectors and proteins or protein-protein complexes, investigated in this work. The restriction sites employed for molecular cloning and the tags encoded at the target N-termini (highlighted in red in the Recombinant vector column) are also reported.

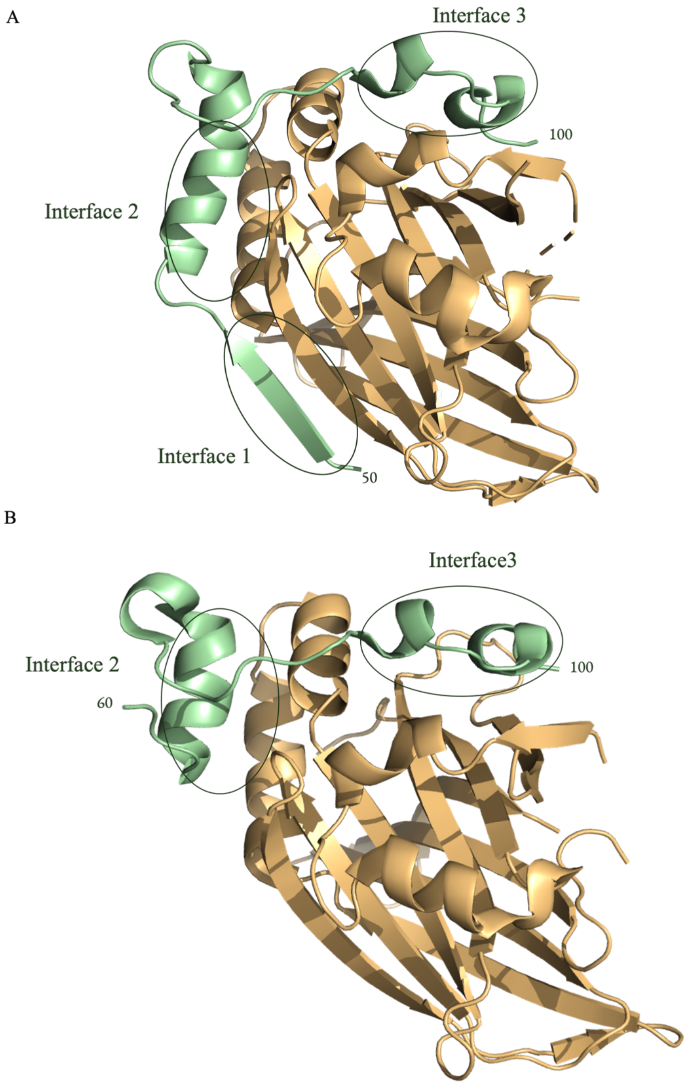


Figure 7.6: Cartoon representation of the (**A**) hYAP:hTEAD1 complex (PDB id 3KYS [161], hYAP and hTEAD1 are in green and gold, respectively); (**B**) hYAP:hTEAD4 complex (PDB id 5OAQ [156], hYAP and hTEAD1 are in green and gold, respectively).The interacting interfaces are highlighted.

Chapter 8

Materials and Methods

8.1 Generation of expression plasmids and production of target proteins and protein complexes

8.1.1 Generation of hTEAD4 YBD expression vectors

The synthetic sequence encoding for hTEAD4 YBD (engineered from the whole gene sequences *hTEAD4*, Acc. No. NM_003213.4, obtained from NCBI, National Center for Biotechnology Information, GenBank database) cloned into the pGEX4-T1 vector within *BamHI/XhoI* restriction sites (*Table 7.1*) was purchased from GenScript USA. This vector also includes the sequence encoding for Glutathione-S-Transferase (GST) that is expressed at the N-terminus of the target protein, obtaining the GST-hTEAD4 YBD fusion protein (*Figure 8.1*). PCR was carried out using the Platinum^{TM} SuperFi^{TM} DNA polymerase (ThermoFisher Scientific^{TM}), in a 50 μl-reaction mixture containing 1 x Superfi^{TM} buffer (ThermoFisher Scientific^{TM}), 0.2 mM dNTPs, 0.5 μM of each primer, 10 ng template DNA and 0.01 unit of DNA polymerase. PCRs were done with an initial denaturation step (30 sec, 98 °C), followed by 30 cycles of denaturation (10 sec, 98 °C), annealing (10 sec, 58 °C), and extension (3 min, 72 °C), and by a final extension step (5 min, 72 °C). The PCR product purified using the NucleoSpin PCR Clean-up Kit (Machery-Nagel), and the vector pET15b were double digested with restriction endonucleases NdeI e BamHI (ThermoFisher Scientific^{TM}), according to the manufacturer's protocol. Following electrophoresis, both DNA fragments were purified from the 1% agarose gel using the E.Z.N.A Gel Extraction Kit (Omega Bio-Tech). Ligation reaction was carried out overnight at 20 °C using 1 unit of T4 DNA ligase (ThermoFisher Scientific^{TM}), 60 ng vector DNA and 36 ng insert DNA (molar ratio insert DNA-vector 5:1). Thereafter, 10 μL of the ligation mixture (pET15b – hTEAD4 YBD plasmid) were then used to heat-shock transform chemically competent *E. coli* TOP-10 cells. Positive transformants were then selected on LB agar plates added by ampicillin (100

Gene Target	Primer	Sequence 5'–3'
hTEAD4 YBD	Fw: hTEAD4-NdeI	GGGAATTCCATATGCGTAGCGTGGCGAGCAGCA
	Rv: hTEAD4-BamHI	CGCGGATCCTTATTCTTTAACCAGACGATAAATG

Table 8.1: Sequences of forward (Fw) and reverse (Rv) primers used for PCR on *hTEAD4 YBD*. The forward and the reverse primers contain the NdeI (CATATG) and the BamHI (GGATCC) restriction sites, respectively (evidenced in red).

mg L^{-1}) and some of them were used to inoculate overnight cultures in LB medium (added by ampicillin 100 mg L^{-1}). After 16 h of incubation at 37 °C, cells were broken, and the ligated plasmids were purified using the E.Z.N.A. Endo-free Plasmid miniprep Kit I (according to the manufacturer's instructions). The ligated recombinant vector was confirmed by restriction digestion analysis with NdeI and BamHI enzymes (ThermoFisher Scientific$^{\text{TM}}$) and then by DNA sequencing (ATAC sequencing, Eurofins Genomics Italy).

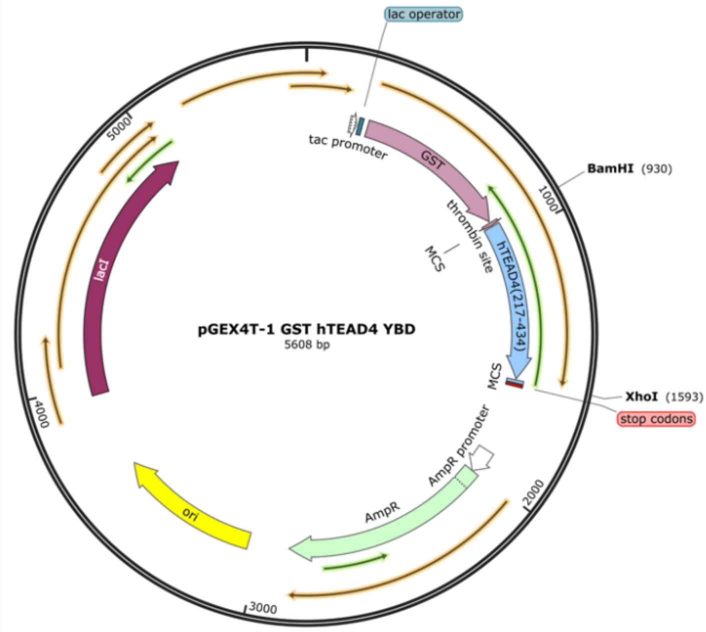


Figure 8.1: Map of the pGEX4T-1 - *hTEAD4 YBD* plasmid vector, including the gene encoding for hTEAD4 YBD cloned within the *BamHI* and *XhoI* restriction sites. The target gene is inserted downstream the sequence encoding for the GST and the thrombin site. AmpR: ampicillin resistance gene; lacI: lactose operon repressor; ori: origin of replication.

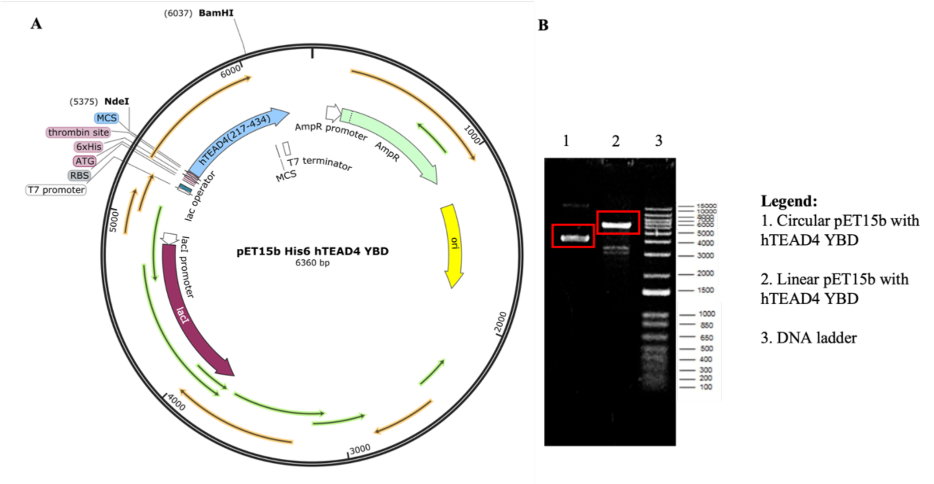


Figure 8.2: (**A**) Map of the final pET15b plasmid vector with *hTEAD4 YBD* sequence. The restriction enzymes NdeI and BamHI have been used for PCR-based cloning.The His^6 tag and thrombin site are reported in the map. AmpR: Kanamycin resistance gene; lacI: lactose operon repressor; ori: origin of replication (**B**) DNA electrophoresis analysis of cloning product. In lanes 1 pET15b His^{6}hTEAD4 YBD (6.360 bp) in circular form; in lane 2 the linearized plasmid which was digested with BamHI enzyme. DNA ladder is reported in lane 5.

8.1.2 *hTEAD4 YBD* expression and purification

The constructs pGEX4-T1-GST-hTEAD4 YBD (*Table 7.1*), allowing the expression of the target protein fused with a thrombin-cleavable GST tag, was heat-shock transformed in chemically competent BL21(DE3) *E. coli* cells. The production and purification of the fusion protein was performed according to the protocols reported by *Bum-Erden et al* [180], with minor modifications. Briefly, cells were grown at 37 °C in LB medium supplemented with ampicillin (100 mg L^{-1}). When optical density at 600 nm (OD$_{600nm}$) reached the value of 0.6-0.8, protein expression was induced using IPTG 0.5 mM; thus the culture was cooled down to 18 °C and maintained under vigorous aeration for other 40 h. Cells were then harvested by centrifugation (3500 g, 8 min, 8 °C) and resuspended in buffer A (PBS buffer, pH 7.3 and 1 mM DTT), supplemented with lysozyme (0.5 mg/mL) and 0.2 mM PMSF. After 1 hour of incubation on ice, cells were disrupted by sonication. The supernatant of the resulting crude extract was collected by centrifugation (13500 g, 1 h, 8 °C) and further purified by glutathione-affinity chromatography on a GST-Trap FF 5 mL column (GE

Healthcare), applied to an Äkta Purifier 10 system. The target protein was eluted in a single step by using an eluting solution composed of 10 mM glutathione, 1 mM dithiothreitol (DTT), and 50 mM Tris HCl pH 8. Fractions containing the protein were identified by SDS-PAGE, pooled and dialyzed in buffer A at 4 °C (membrane cutoff 14 kDa). The GST-tag cleavage was attempted by adding thrombin (13 units mg^{-1} target protein) directly inside the dialysis bag, but even after 72 h of incubation only a small amount of mature protein was produced, as observed by SDS-PAGE analysis. The resulting sample was passed to a second stage of affinity chromatography (GST-Trap FF 5 mL column) and the mature protein was eluted as unbound fraction. The purity of the resulting hTEAD4 YBD was verified by SDS-PAGE analysis. The pET-15b-hTEAD4 YBD expression vector, which includes a thrombin-cleavable His^6-tag added at the N-terminus of the target protein (*Table 7.1*), was used to heat-shock transform *E. coli* ArcticExpress (DE3) cells. The bacterial culture was grown at 30 °C in 1 L of ZYP-5052 auto-induction medium supplemented with 100 mg L^{-1} of ampicillin. When the OD_{600nm} reached the value of 1, the culture was cooled down to 12 °C and cell growth was continued for 60 h, keeping a vigorous aeration. Cells were harvested by centrifugation (3500 g, 20 min, 8 °C) and resuspended in buffer A (25 mM TRIS pH 8.0 and 150 mM NaCl), supplemented with lysozyme (0.5 mg/mL), 0.2 mM PMSF and 20 mM imidazole and disrupted by sonication after 1 hour of incubation on ice. The supernatant of the resulting crude extract was collected by centrifugation (13500 g, 1h, 8 °C) and further purified by nickel-affinity chromatography on a HisTrap FF 5 mL column (GE Healthcare). The target protein was eluted using a three-step gradient protocol, by applying a 250 mM imidazole concentration in buffer A. Fractions containing the target protein were identified by SDS-PAGE, pooled and dialyzed in buffer A at 4 °C (membrane cutoff 14 kDa). The His^6-tag cleavage was performed during the dialysis, by adding thrombin (5 units mg^{-1} target protein) directly inside the dialysis bag. As verified by SDS-PAGE analysis, the complete His^6-tag cleavage was achieved overnight. The resulting sample was applied to a second stage of nickel-affinity chromatography (HisTrap FF 5 mL column) and the mature protein was eluted as weakly-bound fraction. The high purity (estimated as >98%) of the mature hTEAD4 YBD protein sample was verified by SDS-PAGE analysis and mass spectrometry. Liquid Chromatography-Mass Spectrometry, or LC-MS, analyses were performed by the research group of Prof Maria Paola Costi at the University of Modena and Reggio Emilia. The tag removal was confirmed by Western blot (WB) analysis using (using the monoclonal anti-6x His-tag – HRP conjugated, Sigma Aldrich).

8.1.3 Generation of pRSFDuet1 - hYAP TBD - hTEAD4 YBD expression vectors encoding for the S, M and L complexes

The expression vector pRSFDuet1 - *hYAP TBD L fragment* - *hTEAD4 YBD* (*Table 7.1*) was obtained from GenScript USA. This expression vector incudes the genes encoding for *hYAP TBD L fragment and hTEAD4 YBD* in the MCS1 (restriction sites *BamHI/ HindIII*) and MCS2 (restriction sites *NdeI/XhoI*), respectively. In the MCS1, the target gene is preceded by the sequences encoding for a thrombin cleavable His^6-tag, thus the hYAP TBD is expressed as His^6-tagged protein (*Figure 8.3*). This vector was used as template to generate the constructs for the S and M complexes (*Table 7.1*). To obtain the gene sequences for the hYAP S and M fragments, the stop codon TAA was introduced at the position encoding for amino acid 121 and 141, respectively, of the *hYAP TBD L fragment* gene. Site-directed mutagenesis was performed by PCR using partially overlapping primers [107], purchased by Eurofins Genomics Italy (primer sequences are reported in *Table 8.2*). Each PCR was carried out in a 50 μL reaction mixture including 10 ng of DNA template (pRSFDuet1 - *hYAP1 TBD L fragment* - *hTEAD4 YBD*), 1 U of Platinum^{TM} SuperFi^{TM} DNA polymerase, 1 x Superfi^{TM} buffer, 1 x GC enhancer (ThermoFisher Scientific^{TM}), 0.2 mM dNTPs, 0.5 μM of each primer. PCRs were done with an initial denaturation step (30 sec, 98 °C), followed by 30 cycles of denaturation (10 sec, 98 °C), annealing (30 sec, 68 °C or 63 °C, for the S and M fragments, respectively), and extension (2 min, 72 °C), and by a final extension step (5 min, 72 °C). Both PCR products were digested with the Fast Digest DpnI enzyme (ThermoFisher Scientific^{TM}), according to the manufacturer protocol. Afterwards, 10 μL of each reaction mixture were used to heat-shock transform *E. coli* TOP 10 cells. Positive transformants were then selected on LB agar plates added by kanamycin (50 mg L^{-1}) and some of them were used to inoculate overnight cultures in LB medium (added by kanamycin 50 mg L^{-1}). After 16 h of incubation at 37 °C, cells were broken, and the plasmids were purified using the E.Z.N.A. Endo-free Plasmid miniprep Kit I (according to the manufacturer's instructions). The presence of the mutations was confirmed by DNA sequencing (ATAC sequencing, Eurofins Genomics Italy).

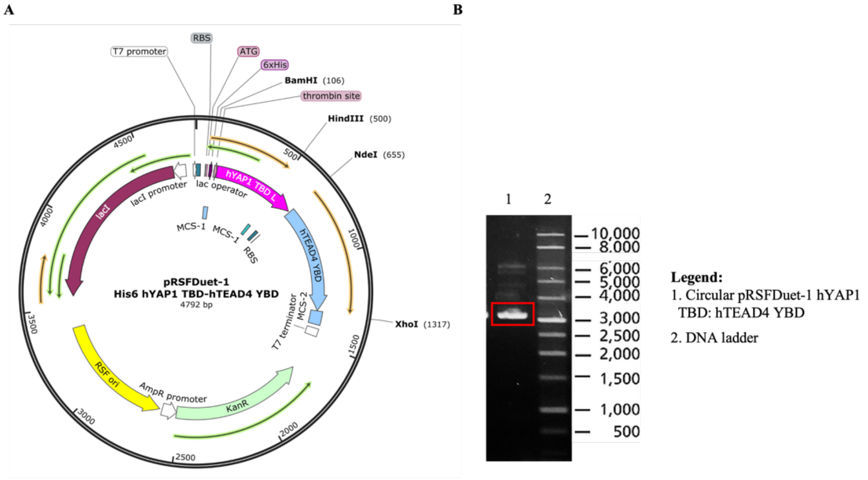


Figure 8.3: (**A**) Map of the pRSFDuet-1 plasmid vector with the sequence encoding for hYAP1 TBD L fragment cloned in MCS1 within BamHI/HindIII restriction enzymes, whereas the sequence encoding for hTEAD4 YBD has been cloned into the MCS2 using NdeI and XhoI enzymes. The His^6 tag and thrombin site are reported in the map. KanR: Kanamycin resistance gene; lacI: lactose operon repressor; RSF ori: origin of replication (**B**) DNA electrophoresis analysis of plasmid on 0.8% agarose gel. In lanes 1 the pRSFDuet-1 hYAP1 TBD L:hTEAD4 YBD (4.792 bp) in circular form. DNA ladder is reported in lane 2.

8.1.4 Co-expression and co-purification of the hYAP TBD:hTEAD4 YBD S, M and L complexes

hTEAD4 YBD and hYAP TBD S, M, L complexes (summarized in *Table 7.1*) were co-expressed and co-purified by means of a single protocol. The expression vector of each complex was used to heat-shock transform *E. coli* BL21(DE3)

GeneTarget	Primer	Sequence5' - 3'	Product
hYAP1 TBD L	Fw: hYAP1 L-STOP 121	GACCCCGTAACATGTGCGTGCGCACAGCAGC	hYAP1 TBD S
	Rv: hYAP1 L-STOP 121	GCACATGTTACGGGGTCAGCGCACCCGCG	
hYAP1 TBD L	Fw: hYAP1 L-STOP 141	GCCCGGGTTAACTGACCCCGACCGGTGTG	hYAP1 TBD M
	Rv: hYAP1 L-STOP 141	GGGTCAGTTAACCCGGGCTAACCGCACCC	

Table 8.2: Sequences of Forward (Fw) and Reverse (Rv) primers used for mutagenic PCR, to introduce a stop codon (TAA reported in red) at the position encoding for amino acid 121 (S fragment) and 141 (M fragment) of the hYAP TBD L fragment gene.

cells. The production of the protein complexes was performed by culturing transformed bacteria in Super Broth (SB) medium added by 50 mg L^{-1} kanamycin. When OD$_{600nm}$ reached values of 0.6-0.8, protein over-expression was induced by adding 0.2 mM IPTG and incubating cells for at 20 °C for 48 h, under vigorous aeration. Cells, harvested by centrifugation, were resuspended in buffer A (250 mM NaCl and 50 mM Tris-HCl, pH 8) added by 20 mM imidazole, lysozyme (0.5 mg mL^{-1}) and 0.2 mM PMSF, and disrupted by sonication after 1-h incubation on ice. The soluble cellular fraction was clarified by centrifugation (13500 x g, 1 h, 8 °C) and the protein complex was purified by taking advantage of the N-terminal His^{6}-tag on the hYAP TBD fragment by means of nickel affinity chromatography (HisTrap FF 5 mL column). The purification protocol relied on a three-step concentration gradient of imidazole (40-250-500 mM in the buffer A), resulting in the elution of His^{6}-tag complexes at imidazole concentrations ranging from 40 to 250 mM. Fractions containing the protein complexes were identified by SDS-PAGE, pooled and dialyzed overnight against buffer A at 4 °C (membrane cutoff 3 kDa). The samples including the His^{6}-tag complexes were then concentrated and applied to an HiLoad 16/600 Superdex 75pg gel filtration column (GE Healthcare) equilibrated in buffer A. Eluted fractions containing the target protein complexes were then concentrated to final concentration of 20 mg mL^{-1} for each complex, using Pierce^{TM} Protein Concentrator PES (cut-off 10 kDa, Thermo Fisher Scientific). The high purity of the samples was confirmed by SDS-PAGE, mass spectrometry, and Western blot analyses (using the monoclonal anti-6x His-tag – HRP conjugated, Sigma Aldrich).

8.1.5 Generation of expression vectors for full-length hYAP and for its complex with hTEAD4 YBD.

The synthetic sequence encoding for *hYAP* cloned into the pMAL-p5X vector within *BamHI* and *HindIII* restriction sites was purchased from GenScript (*Table 7.1*). This vector also includes the sequence encoding for the Maltose Binding Protein (MBP) which is expressed at the N-terminus of the target protein, obtaining the MBP-hYAP fusion protein. Furthermore, the gene coding sequence for the thrombin cleavage site was also added between the two proteins to allow the tag removal by means of this protease and to including it in the following sub-cloning in the pQE-80L vector. Indeed, since this vector encodes for a different affinity tag, His^6-tag, but it does not encode for a protease cleavage site, we decided to include this additional coding sequence before our target gene, to allow yielding the mature hYAP protein. On this purpose, the *hYAP* gene (preceded by the thrombin-cleavage-site coding sequence) was then sub-cloned in the vector pQE80-L (Novagen), within the *BamHI* and *HindIII* restriction sites (*Table 7.1* and *Figure 8.4*). The *hYAP* gene (preceded by the thrombin-cleavage-site coding sequence) was also sub-cloned in the MCS 1 (within the *BamHI* and *HindIII* restriction sites) of the pRSFDuet-1 - *hYAP TBD L fragment* - *hTEAD4 YBD* plasmid (*Table 7.1*), thus replacing the *hYAP TBD L fragment* gene. The resulting expression vector, pRSFDuet-1 - *hYAP* - *hTEAD4 YBD* allows the production of the hTEAD4 YBD in complex with full length hYAP (*Table 7.1* and *Figure 8.5*). For each sub-cloning, the *hYAP* gene (preceded by the thrombin-cleavage-site coding sequence) and the plasmids were obtained by the double digestion with the restriction enzymes BamHI and HindIII, performed at 37 °C for 16 h. Ligation reactions were carried out overnight at 20 °C using 1 unit of T4 DNA ligase. For the *hYAP* gene cloning into pQE-80L, 60 ng of vector DNA and 85 ng of insert DNA (molar ratio insert DNA-vector 5:1) have been used. On the other hand, 40 ng of vector DNA and 28 ng of insert DNA (molar ratio insert DNA-vector 2:1) were used to obtain pRSFDuet-1 - *hYAP* - *hTEAD4 YBD* plasmid (*Figure 8.3*). Then, 10 μL of the ligation mixtures were then used to heat-shock transform chemically competent

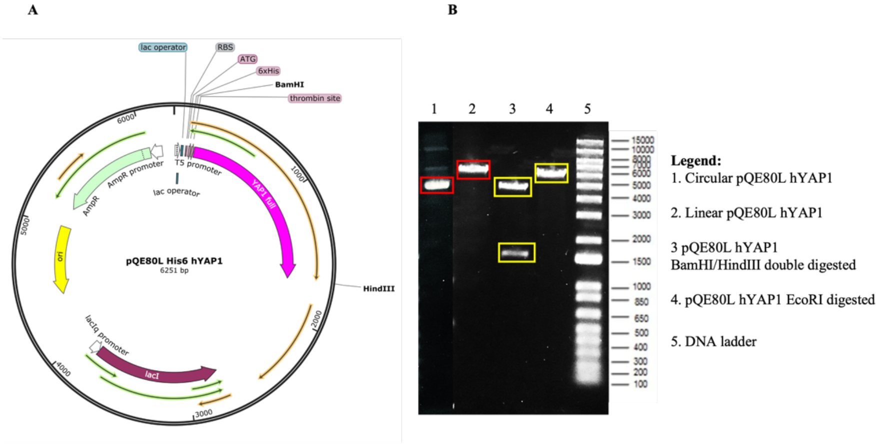


Figure 8.4: (**A**) Map of the final pQE80L plasmid vector with *hYAP1* gene cloned into BamHI/HindIII restriction sites. The His^6 tag and thrombin site are reported in the map. AmpR: Ampicillin resistance gene; lacI: lactose operon repressor; ori: origin of replication (**B**) DNA electrophoresis analysis of cloning product on 1.2% agarose gel. In lanes 1 and 2 the pQE80L hYAP1 (6.251 bp) in circular and linear form, respectively. In lane 3 the restriction digestion of recombinant vector with enzymes BamHI and HindIII resulted in a vector fragment with (4.702 bp) and hYAP1 gene (1.549 bp). DNA ladder is reported in lane 5.

E. coli TOP-10 cells. Positive transformants were then selected on LB agar plates added by either ampicillin (100 mg L^{-1}) or kanamycin (50 mg L^{-1}), for the pQE-80L – *hYAP* and pRSFDuet-1 - *hYAP* - *hTEAD4 YBD* plasmid, respectively. Some transformants were singularly used to inoculate LB overnight cultures, afterwards used to extract the ligated expression plasmids using the E.Z.N.A. Endo-free Plasmid miniprep Kit I (according to the manufacturer's instructions). The ligated recombinant vectors were confirmed by restriction digestion analysis with BamHI and HindIII enzymes and then by DNA sequencing (ATAC sequencing, Eurofins Genomics Italy).

8.1.6 Production of full length hYAP1

The production of both MBP-hYAP and His^6-hYAP was attempted in different *E. coli* strains, ArcticExpress (DE3), BL21(DE3) and BL21(DE3) pLysS, testing also different culture media (LB, SB, and ZYP5052), and protein over-production conditions (incubation temperatures of 28 °C and 18 °C and times of 24 h and 48 h, and IPTG concentrations of 0.1 mM and 1 mM, in LB and SB media). Regardless the production conditions, both MBP-hYAP and His^6-hYAP resulted

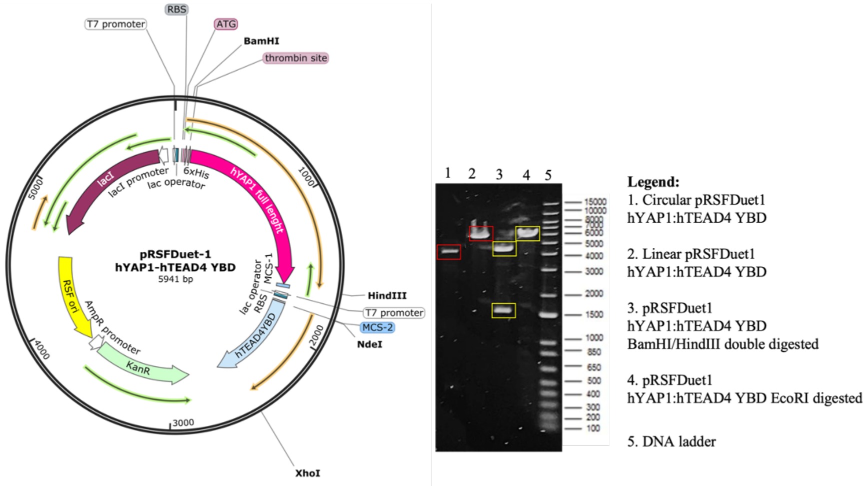


Figure 8.5: (**A**) Map of the final pRSFDuet-1 plasmid vector with *hYAP1* gene cloned in MCS1 within BamHI/HindIII restriction enzymes; the sequence encoding for hTEAD4 YBD has been cloned into the MCS2 using NdeI and XhoI enzymes. The His^6-tag and thrombin site are reported in the map. KanR: Kanamycin resistance gene; lacI: lactose operon repressor; RSF ori: origin of replication (**B**) DNA electrophoresis analysis of cloning product. In lanes 1 and 2 the pRSFDuet-1 hYAP1:hTEAD4 YBD (5.941 bp) in circular and linear form, respectively. In lane 3 the restriction digestion of recombinant vector with enzymes BamHI and HindIII resulted in a vector fragment with hTEAD4 YBD sequence (4.392 bp)and hYAP1 gene (1.549 bp). DNA ladder is reported in lane 5.

poorly soluble in *E. coli*, being mainly localized in the insoluble cellular fractions (inclusion bodies). These restrictions in the production of the target protein did not allowed us to obtain the isolated hYAP and its further characterization.

8.1.7 Co-expression and co-purification of the hYAP1:hTEAD4 YBD complex (also named as XL complex)

The production of His^6-tag full length hYAP in complex with hTEAD4 YBD (XL complex) was performed in the *E. coli* ArcticExpress (DE3) system, using the ZYP-5052 autoinduction medium (supplemented with 50 mg L^{-1} kanamycin). Bacterial cells were grown at 30 °C until the OD$_{600nm}$ reached values of ≈1, then the culture was cooled down at 12 °C and incubated for additional 60 h, maintaining the vigorous aeration. Cells, harvested by centrifugation (3500 g, 20 min, 8 °C), were resuspended in buffer A (250 mM

NaCl, 50 mM Tris, pH 8), supplemented with 20 mM imidazole, lysozyme (0.5 mg mL^{-1}) and PMSF (0.2 mM), and disrupted by sonication after 1 h incubation on ice. The supernatant of the resulting crude extract was collected by centrifugation (12000 g, 1 h, 8 °C) and further purified by nickel-affinity chromatography (HisTrap FF 5 mL column). The purification protocol relied on a three-step concentration gradient of imidazole (40-250-500 mM in buffer A). The His^6-tag complex was eluted by applying imidazole concentrations of 40 mM and 250 mM and the collected fractions (identified by SDS-PAGE) were pooled and dialyzed overnight in buffer A at 4 °C (membrane cutoff 14 kDa). The protein sample was then concentrated and further purified by gel filtration, using a HiLoad 16/600 Superdex 200pg coloumn (GE Healthcare) equilibrated in buffer A. Four main peaks, having retention volumes equal or lower than the one expected for the protein complex (estimated on its collective molecular weight), were separated from the column, and further characterized by SDS-PAGE analysis. Bands corresponding to His^6-tag hYAP and hTEAD4 YBD of the target complex were identified only in two main peaks, named F3 and F5 having retention volumes compatible with the heterodimer and an aggregate of two heterodimers, respectively (as also supported by native PAGE analysis, see below in **section 9.3.2**. The presence of His^6-tag hYAP was verified in both fractions by Western blot analysis, (using the monoclonal anti-6x His-tag – HRP conjugated, Sigma Aldrich). The matching of the SDS-PAGE protein bands with His^6-tag hYAP and hTEAD4 YBD, forming the target complex, was also verified by peptide mass fingerprinting analysis, performed on service in Toscana Life Sciences by Dr. Laura Salvini.

8.1.8 Native PAGE and Mass Spectrometry analyses

The purified hYAP TBD:hTEAD4 YBD L and hYAP:hTEAD4 YBD XL complexes were analyzed by native PAGE, using NativePAGE^{TM} 4 -16% Bis-Tris Gels (Thermo Fisher Scientific^{TM}), according to the manufacturer protocol. Briefly, samples (including a total protein amount of ≈1 μg) were run at 4 °C by applying a constant current of 150 V for 60 min, then increased to

250 V for additional 60 min. The NativePAGE^{TM} Running Buffer was used as anode buffer, whereas the cathode buffer was the Light Blue Cathode Buffer (prepared by adding 0.002 % Coomassie G-250 to the NativePAGE^{TM} Running Buffer). The native gels were stained with the SimplyBlue Safe Stain Coomassie G-250 (Thermo Fisher Scientific^{TM}) and detection of His^{6}-tag-complexes was confirmed by Western blot analysis (using the monoclonal anti-6x His-tag – HRP conjugated, Sigma Aldrich). Native mass-spectrometry analysis on hYAP TBD:hTEAD4 YBD L was performed on service in Toscana Life Sciences by Dr. Laura Salvini.

8.1.9 Circular dichroism (CD) and thermal denaturation analyses

The circular dichroism (CD) spectra of the mature hTEAD4 YBD and of the four complexes were measured on a Jasco J-815 spectropolarimeter at room temperature, using a 0.1 cm quartz cuvette and the wavelength range 250-200 nm. All protein sample had a concentration of 20 μM in 25 mM Tris pH 8 and 150 mM NaCl. Each protein sample was also used to perform thermal denaturation analysis by monitoring the CD signal at 220 nm in the temperature range 25 °C - 100 °C (temperature increase rate of 2 °C min^{-1}). All measurements were performed in triplicates. The midpoint of the unfolding transition (T_m) was determined for each thermal denaturation profile through the software GraphPad Prism 7 in non-linear regression procedure applying the Boltzmann sigmoidal function.

8.1.10 Crystallization of hTEAD YBD and its complexes with the hYAP TBD fragments and full-length hYAP

Crystallization trials on hTEAD YBD and its complexes with the hYAP TBD fragments were performed using the commercially available kits JBScreen Basic (JBSB) 1–4 and Classic (JBSC) 6 from Jena Bioscience (Jena, Germany), and PEG/Ion 1 and 2, Index, Crystal Screen 1 and 2, and Grid screen Ammonium

Sulfate from Hampton Research (Aliso Viejo, California, USA). On the other hand, the lower amount of sample obtained for the heterodimer hYAP:hTEAD YBD limited the screens to PEG/Ion 1 and 2. Drops, consisting of 1 μL protein solution (20 mg mL^{-1} for hTEAD YBD and its complexes with the hYAP TBD fragments and 12 mg mL^{-1} for the heterodimer hYAP:hTEAD YBD) and 1 μL precipitant, were equilibrated over 200 μL reservoir. Crystallization screens were performed using the sitting-drop vapor diffusion technique [109], both at 8 °C and 20 °C (only the latter temperature was attempted for the heterodimer hYAP:hTEAD YBD). To date, no crystal formation was observed for hTEAD YBD and for the heterodimer hYAP:hTEAD YBD. On the other hand, crystals of all S, M, and L complexes grew using different precipitant solutions (*Tables* **??**) both at 8 °C and 20 °C. Unfortunately, all crystals showed poor diffraction patterns in preliminary X-ray crystallographic analyses. Thus, to improve the crystal ordering/quality, the micro-seeding crystallization technique was applied [181]. For each successful condition, crystals were crushed and used to prepare the mother seeding solution. Seeding solutions at dilutions 1:1000 and 1:10000 were then prepared by following dilutions of the mother solution with each precipitant. The 1:1000 and 1:10000 seeding solutions were used for streak-seeding crystallization experiments, as summarized in *Table 9.12*. Crystals, grown within few days, were subjected to X-ray crystallographic analysis to assess improvements in their diffraction patterns. All crystals were cryoprotected using a solution composed of precipitant enriched by 20% v/v ethylene glycol and flash frozen in liquid nitrogen. For the complex S, the best diffracting crystals were obtained using 0.2 M calcium chloride dihydrate, 15% PEG 3350 as precipitant solution and 5% v/v of 1:1000 seeding solution.

8.1.11 Data collection, structure solution and refinement

A large amount (>200) of protein crystals were screened for diffraction using synchrotron radiation at the Diamond Light Source (DLS, Didcot, UK) beamlines I03 and I04 equipped with Eiger2 XE 16M detector. Despite this large crystal screening campaign, only few crystals of the S complex, grown using the seeding technique, provided diffraction patterns at resolution <3 Å. Full data sets were collected on these crystals, integrated with XDS [110] and scaled through the CCP4 suite [113] programs SCALA and AIMLESS [111, 112]. After these initial phases of data processing, the strongest and better-quality dataset was selected and used for proceeding in structure solution and refinement. Data collection and processing statistics are reported in *Table 8.3.* By the Matthews-Cell content analysis [182], 12 heterodimers were estimated to populate the crystal ASU. The structure was solved by molecular replacement using the software Molrep [114] and one heterodimer of hYAP TBD:hTEAD4 YBD (PDB id 6GE3 [163], the YAP fragments used to obtain this complex includes only residues 60-99) as searching model. The presence of Non-Crystallographic Symmetry (NCS) [183] was assessed by the calculation of the Native Patterson Map through the FFT Patterson program [184]. The structural model of the S complex was refined using REFMAC5 [115], and the molecular graphic software Coot [116] was used for electron density inspection, model rebuilding and to add water molecules. The final model was inspected manually and checked with the programs Coot [116] and Procheck [117]. Final refinement statistics are reported in *Table 8.4.* Figures were generated using the molecular graphic softwares PyMol [118] and CCP4mg [185].

8.1.12 Preliminary Bio-SAXS and CryoEM studies of the hYAP:hTEAD4 YBD complex (or XL complex)

Preliminary Size Exclusion Chromatography–Biological Small Angle X-ray Scattering (SEC-Bio SAXS) data were collected at European Synchrotron Radiation Facility (ESRF) on the beamline BM29, equipped with a Pilatus3 2M detector. This preliminary analysis was performed in collaboration with Dr Dritan Siliqi (Institute of Crystallography- CNR, Bari, Italy) and Dr Giancarlo Tria (Department of Chemistry "Ugo Shiff", University of Florence, Italy). The experiments were carried out on both the samples of both the heterodimer His^6-hYAP:hTEAD4 YBD (XL complex) and the aggregate of two heterodimers of His^6-hYAP:hTEAD4 YBD (2x-XL complex). Different protein samples of XL

	His^6-hYAP TBD:hTEAD4 YBD S complex
Diffraction source	I04 (DLS)
Wavelength (Å)	0.9795
Temperature (K)	100
Detector	Eiger2 XE 16M
Crystal-detector distance (mm)	385.5
Exposure time per image (s)	0.15
Space group	P$3_1$21
No. of heterodimers in the ASU	12 (12 hYAP TBD S fragment and 12 hTEAD4 YBD
a = b, c (Å)	164.58, 258.61
Resolution range (Å)	73.76-2.50 (2.54-2.50)
Total no. of reflections	1092787 (53844)
No. of unique reflections	140121 (6836)
Completeness (%)	100.0 (100.0)
Redundancy	7.8 (7.9)
$\langle I/\sigma(I) \rangle$	19.3 (2.3)
CC1/2	0.99 (0.75)
R_{meas}	0.072(1.013)
Overall B factor from Wilson plot (Å^2)	56.2

Table 8.3: Data collection and processing statistics. Values in parentheses are for the highest resolution shell.

	His^6-hYAP TBD:hTEAD4 YBD S Complex
Resolution range (Å)	73.76-2.50 (2.54-2.50)
Completeness (%)	100 (100)
No. of reflections, working set	133152 (9737)
No. of reflections, test set	6911 (503)
R_{cryst}	20.47 (28.5)
R_{free}	25.72 (33.9)
No of non-H atoms	
Protein	23615
Myristic acid	175
Other (triethylene glycol)	10
Water	501
Total	24301
R.m.s. deviations	
Bonds (Å)	0.006
Angles (°)	1.554
Average B factors (Å^2)	57.82
Estimate error on coordinates based on	
R value (Å)	0.428
Ramachandran plot	
Most favored (%)	92.1
Allowed (%)	7.9

Table 8.4: Refinement statistics. Values in parentheses are for the highest resolution shell.

and 2x-XL complexes at concentrations of 6 mg mL^{-1} and 12 mg mL^{-1} were applied to a KW-403 SEC column, equilibrated with buffer A (25 mM Tris pH 8, 150 mM NaCl), and the eluate was examined in-line. Furthermore, protein samples of XL and 2x-XL complexes at concentrations of 3 mg mL^{-1} and 6 mg mL^{-1} were used for batch measurements with the robotic sample changer. Preliminary cryo-EM analysis on the His6-hYAP1:hTEAD4 YBD XL complex were performed on service at the Florence Center for Electron Nanoscopy (FloCEN), University of Florence in collaboration with Dr Giancarlo Tria and Dr Annalisa Guerri. Samples of both XL and 2x-XL complexes were analyzed at concentrations of 2, 3 and 6 mg mL^{-1}. For each sample, a volume of $\approx$3 μL was applied on glow-discharged Quantifoil Cu300 R1.2/1.3 Cu 300 mesh grids and the excess of liquid was removed by blotting for 2 sec (blot force 0) using filter paper, followed by plunge freezing in liquid ethane. Vitrification was done using a FEI Vitrobot Mark IV at 100 % humidity and 10 °C. Cryo-EM images were then recorded on a Glacios microscope (ThermoFisher Scientific^{TM}), at 200 kV with a Falcon3 detector operating in integration mode, with pixel size of 0.96 Å and a total dose of 53 e^{-} A^{-2} over an exposure time of 2 seconds.

Chapter 9

Results and Discussions

9.1 Studies on hTEAD4 YBD

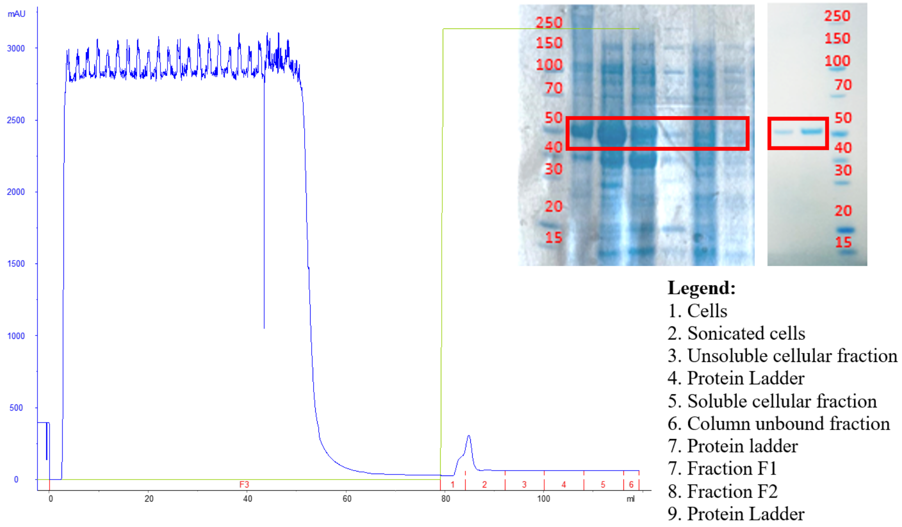
9.1.1 GST-hTEAD4 YBD vs His^6-hTEAD4 YBD

The first attempting to obtain hTEAD4 YBD were made as GST-fusion protein (the whole sequence of the GST-hTEAD4 YBD fusion protein is shown in *Table* **??**) by cloning the target gene in the pGEX4-T1 plasmid (*Figure 8.1*). GST-fusion proteins are broadly used to increase the solubility of a target in the expression system and facilitate their purification by exploiting affinity chromatographic techniques [186]. Indeed former studies reported the production of hTEAD4 YBD in *E. coli* BL21 (DE3) as GST-fusion protein [180]. Despite the application of analogous expression condition, we observed that GST-hTEAD4 YBD was mainly present in the insoluble fraction in *E. coli* cells. Even so, we attempted to isolate the fusion protein present in the soluble fraction, using glutathione-affinity chromatography (chromatograms and SDS-PAGE analyses of GST-hTEAD4 YBD purification are shown in *Figures 8.5 and 9.2*). A very low amount (3 mg) of highly pure GST-hTEAD4 YBD was obtained after the first stage of affinity chromatography (*Figure 9.1 A*). Despite the high protease amount and long incubation time, attempting to obtain the mature hTEAD4-YBD resulted only in very low cleavage yields (as shown in SDS-PAGE analysis reported in *Figure 9.1 B*). Indeed, in the second stage of glutathione-affinity chromatography we covered almost only the uncleaved GST-hTEAD4 YBD, with very low yield, estimated to 1 mg L^{-1} bacterial culture (*Figure 9.2*). The lack of an effective thrombin cut could be reasonably due to the hindrance generated by the fusion proteins that hide the protease recognition site. Given the problems encountered with this construct, we decided to try a different, less hindering affinity tag, producing the target protein as His^6-hTEAD4-YBD. On this purpose we cloned the *hTEAD4 YBD* gene into the pET-15b vector (*Figure 9.7*), allowing the expression of the His^6-hTEAD4 YBD recombinant protein (sequence in *Table 9.1*).

The production trials performed to maximize the yields of His^6-hTEAD4 YBD evidenced that a high amount of the target protein is present in the *E.*

Sample	Studied protein-protein complexes	Sequences
hTEAD4 YBD	His^{6}hTEAD4 (217-434)	(-21)MGSSHHHHHHSSGLVPR↓ GSH (1)MRSVASSKLWMLEFSAFLEQQQDPDTYNKHLFVHIGQSSPSYSDPYLEAVDIRQIYD KFPEKKGGLKDLFERGPSNAFFLVKFWADLNTNIEDEGSSFYGVSSQYESPENMIITCST KVCSFGKQVVEKVETEYARYENGHYSYRIHRSPLCEYMINFIHKLKHLPEKYMMNSV LENFTILQVVTNRDTQETLLCIAYVFEVSASEHGAQHHIYRLVKE(219)
hTEAD4 YBD	GST-hTEAD4 (217-434)	(-220)MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNL PYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDF ETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDA FPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLVPR↓ GS (1)MRSVASSKLWMLEFSAFLEQQQDPDTYNKHLFVHIGQSSPSYSDPYLEAVDIRQIYDK FPEKKGGLKDLFERGPSNAFFLVKFWADLNTNIEDEGSSFYGVSSQYESPENMIITCSTKV CSFGKQVVEKVETEYARYENGHYSYRIHRSPLCEYMINFIHKLKHLPEKYMMNSVLENF TILQVVTNRDTQETLLCIAYVFEVSASEHGAQHHIYRLVKE(219)

Table 9.1: Summary of the investigated GST-hTEAD4 YBD and His^6-hTEAD4 YBD. Amino acids belonging to the GST and His6 tag are colored in green; the thrombin cleavage sites are red, respectively, while the sequences of hTEAD4 YBD is in black. The black arrow indicates the site of cleavage by thrombin protease.



Legend:
1. Cells
2. Sonicated cells
3. Unsoluble cellular fraction
4. Protein Ladder
5. Soluble cellular fraction
6. Column unbound fraction
7. Protein ladder
7. Fraction F1
8. Fraction F2
9. Protein Ladder

Figure 9.1: (**A**) Chromatogram of the first stage of Ni^{2+} affinity chromatography (HisTrap FF 5mL column, GE-Healthcare) performed to purify His^6-hTEAD4 YBD. The UV$_{280\text{nm}}$ monitoring and imidazole concentration are shown as blue and green lines, respectively; the elution fractions are indicated on the x-axis as red marks. (**B**) SDS-PAGE analysis of the fractions collected from affinity chromatography (the protein ladder used to estimate the molecular weight, or MW, is in lane 6).

coli ArcticExpress (DE3) soluble cellular fraction. Thus, we produced the target through an efficient, reliable protocol, developed for heterologous protein expression in this peculiar system, relying on culturing bacteria in the ZYP-5052 autoinduction medium at 30 °C to high OD$_{600\text{nm}}$ values (>1), followed by a 60 h incubation at 12 °C. The purification protocol developed to

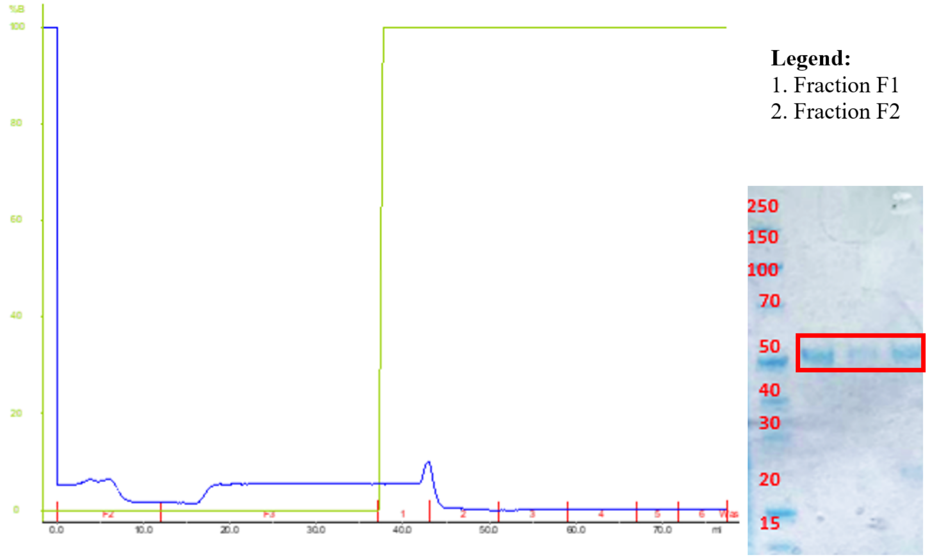


Figure 9.2: (**A**) Chromatogram of the first stage of Ni^{2+} affinity chromatography (HisTrap FF 5mL column, GE-Healthcare) performed to purify His^6-hTEAD4 YBD. The UV_{280nm} monitoring and imidazole concentration are shown as blue and green lines, respectively; the elution fractions are indicated on the x-axis as red marks. (**B**) SDS-PAGE analysis of the fractions collected from affinity chromatography (the protein ladder used for MW estimation is in lane 1).

obtain the mature hTEAD4 YBD relied on a two-stage nickel affinity chromatography. The highly pure His^6-hTEAD4 YBD obtained after the first stage (chromatogram and SDS-PAGE analysis in *Figure 9.3*), was thus cleaved by thrombin protease, and the mature hTEAD4 YBD separated by the second step of affinity chromatography (chromatogram and SDS-PAGE analysis in *Figure 9.4*). The final hTEAD4 YBD yield obtained by these procedures was estimated to 40 mg L^{-1} bacterial culture, reporting a drastic increment in target production with respect to the former attempting with the GST-hTEAD4 YBD fusion protein.

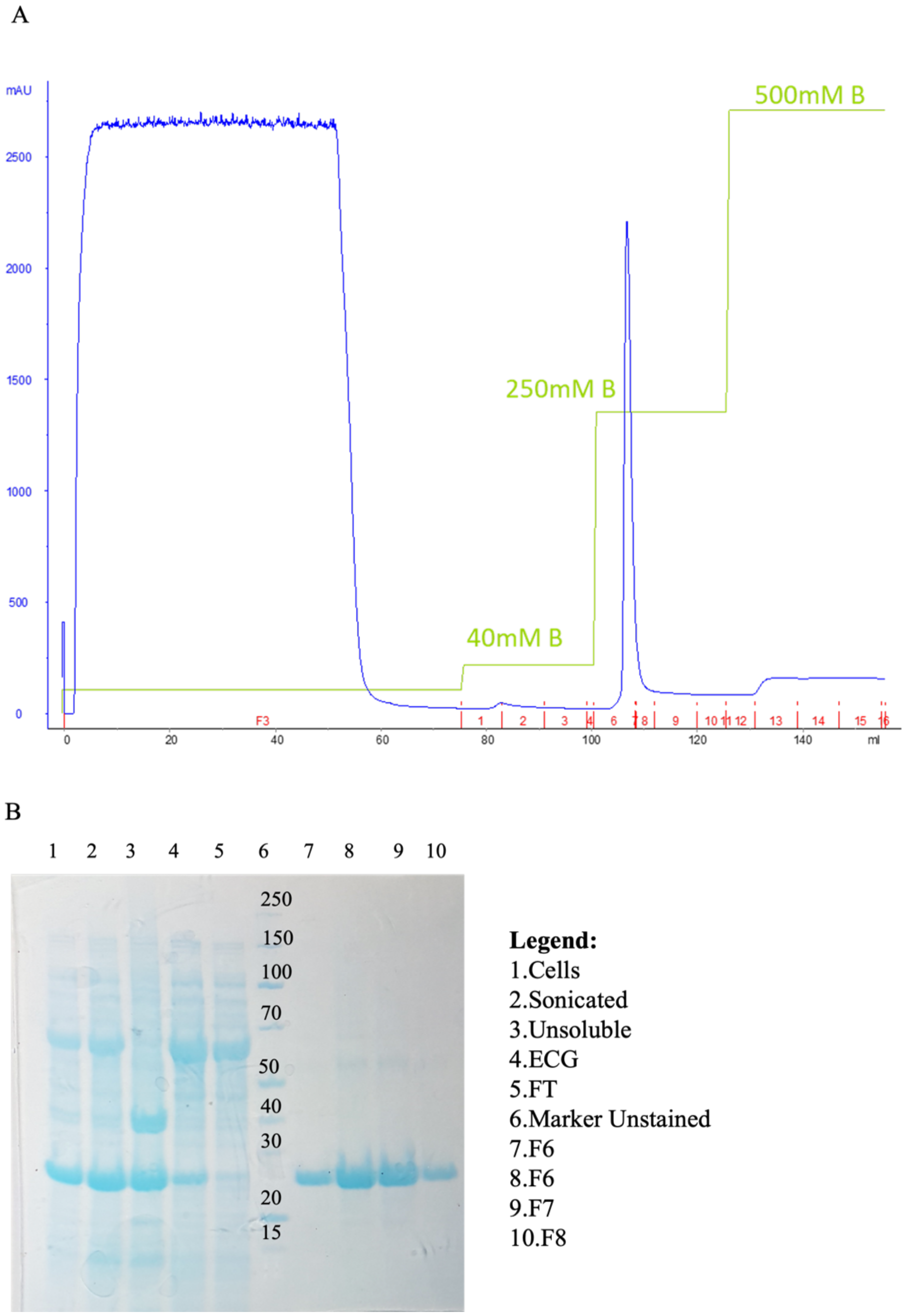


Figure 9.3: (**A**) Chromatogram of the first stage of nickel- affinity chromatography (HisTrap FF 5mL column, GE-Healthcare) performed to purify His^{6}-hTEAD4 YBD. The UV$_{280\mathrm{nm}}$ monitoring and imidazole concentration are shown as blue and green lines, respectively; the elution fractions are indicated on the x-axis as red marks. (**B**) SDS-PAGE analysis of the fractions collected from affinity chromatography (the protein ladder used for MW estimation is in lane 6).

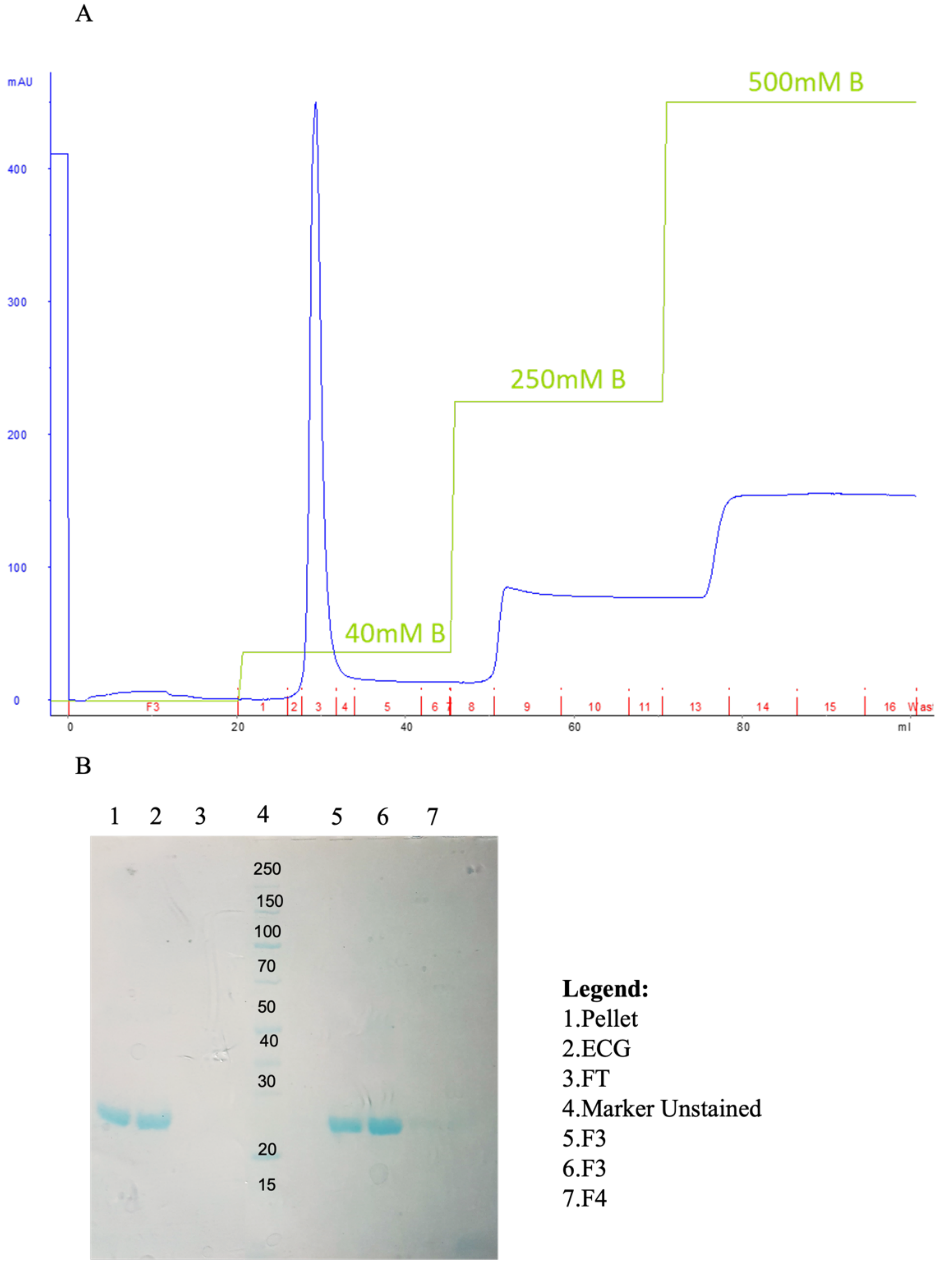
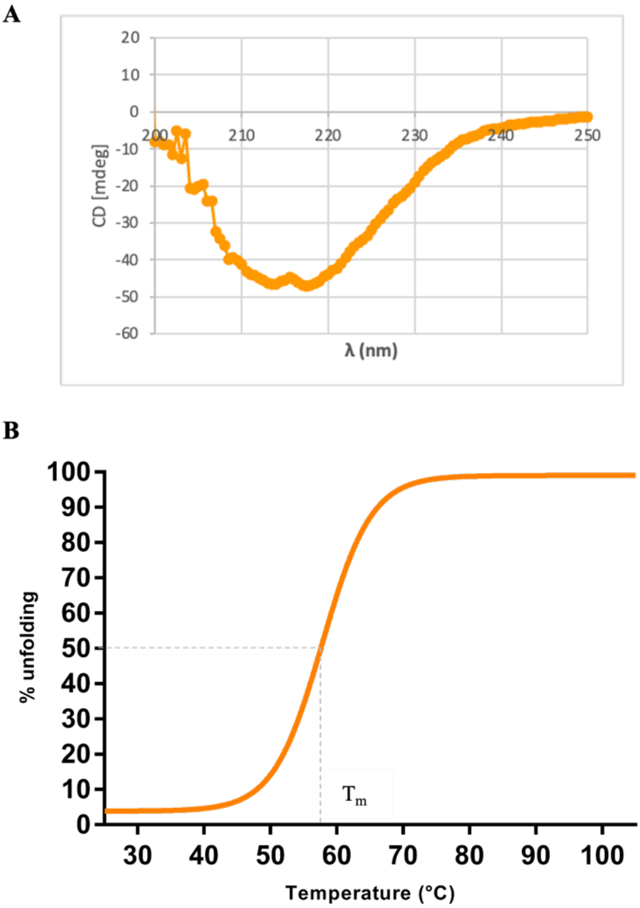


Figure 9.4: (**A**) Chromatogram of the second stage of nickel- affinity chromatography (HisTrap FF 5mL column, GE-Healthcare) performed to purify His^6-hTEAD4 YBD. The UV$_{280nm}$ monitoring and imidazole concentration are shown as blue and green lines, respectively; the elution fractions are indicated on the x-axis as red marks. (**B**) SDS-PAGE analysis of the fractions collected from affinity chromatography (the protein ladder used for MW estimation is in lane 4).

9.1.2 Characterization of recombinant, mature hTEAD4 YBD

The mature hTEAD4 YBD, obtained from the production as His^6-tagged protein, was employed for characterizations through LC-MS, CD and thermal denaturation analyses, and finally for structural investigations. The LC-MS analysis, performed by the research group of Prof Maria Paola Costi at the University of Modena and Reggio Emilia, showed the presence of three main species: hTEAD4 YBD (MW= 25808.6) and two acylated forms attributed as myristoylated hTEAD4 YBD (MW = 26018.9, Myr-hTEAD4 YBD) and palmitoylated hTEAD4 YBD (MW = 26045.9, Palm-hTEAD4 YBD). According to their analyses and in agreement with the literature [156], the non-acylated hTEAD4 YBD constitutes the 25-40% of the sample. Thus, the target protein seems to be mainly modified inside the expression system, resulting mainly present as Myr-hTEAD4 YBD. The hTEADs have a deep hydrophobic central cavity in their YBD that can be populated by myristic or palmitic acid, acylating the conserved cysteine residue (Cys367 for TEAD4) exposed inside the pocket, under physiological conditions [187]. Furthermore, it has been reported that under physiological concentrations of palmitoyl-CoA, TEADs can undergo PAT-independent auto-palmitoylation [157].

The mature hTEAD4 YBD was also used for CD and thermal denaturation analyses, to verify the protein folding and to measure its thermal stability. The CD spectra recorded on hTEAD4 YBD is coherent with a folded protein and with reported structures [156, 163]. The same sample was thus used to characterize the thermal unfolding profile of the protein, by monitoring the far-UV CD signal at 220 nm in the temperature range 25-105 °C (*Figure 9.5*). The T_m value determined by this analysis for hTEAD4 YBD is of 57.82 (± 0.1) °C. Former investigations by fluorescence-based thermal denaturation analysis (FTDA) [156], reported lower T_m values, 43.7 (± 0.3) °C, for the non-acylated hTEAD4 YBD, and higher T_m values, 54.1 (± 0.2) °C, for the acylated protein. Upon acylation with myristic or palmitic acid, the protein stability increases, leading to an increment in the T_m value of $\approx$10 °C [156]. The T_m value determined by us through CD-thermal denaturation analysis correlates with that of the acylated form obtained by FTDA. This data is also in agreement with LC-MS data, showing that in our sample hTEAD4 YBD is mainly present in the myristoylated form. An analogous effect of increment in protein stability upon acylation was also reported for hTEAD2 [157]. Nonetheless, the role of hTEAD acylation has not been fully clarified yet. Although the extensive attempting of protein crystallization we have not been able to obtain crystals of the isolated hTEAD4 YBD. Probably, the presence of a fraction of non-acylated protein in the purified sample has a negative influence on the crystallization process.



Protein sample	Melting Temperature (Tm)°C
hTEAD4-YBD	57.82 ± 0.1

Figure 9.5: (**A**) Circular dichroism spectrum of His^6- hTEAD4 YBD (217-434) protein. (**B**) Thermal unfolding transition curves of hTEAD4 YBD followed by circular dichroism. Measurements were performed in triplicate. Melting temperature (T_m) values determined for the transitions is reported in the table.

9.2 Studies on the hYAP TBD: hTEAD4 YBD S, M and L complexes

9.2.1 Co- expression and co-purification

The protocol previously described to study the interaction between hTEAD4 YBD and hYAP TBD relied on the production of hTEAD4 YBD in *E. coli* cells and the subsequent formation of the complex by its combination with a synthetic peptide including the hYAP TBD fragment 60-99 (syn-hYAP TBD 60-99). On the other hand, in this work, we have established valid and reliable protocols for co-expression and co-purification of hTEAD4 YBD in complex with different hYAP TBD fragments (*Table 7.1*). By using the DUET vector technology, the hTEAD4 YBD was co-expressed with hYAP1 TBD fragments 50-120 (S complex), 50-140 (M complex) and 50-170 (L complex) (*Tables 9.3* and **??**). The constructs or the S and M complexes were generated by introducing stop codons at proper positions of the *hYAP TBD* gene on the expression vector of the complex L through PCR-based site-direct mutagenesis (*Figure 9.8*). All complexes were mainly expressed in the soluble cellular fraction of *E. coli* BL21(DE3) cells, thus we employed this bacterial strain for their production, using a different protocol with respect to that developed for the isolated hTEAD4 YBD. The purification of the complexes, relying on nickel affinity chromatography, took advantage of the His^6-tag present at the N-terminus of hYAP1, whereas the hTEAD4 YBD was expressed as un-tagged protein, thus it was co-purified with hYAP1 through the formation of the heterodimeric complex (the amino acid sequences of all complexes are summarized in *Table 9.2*).

After a single stage of nickel affinity chromatography, highly pure (>98%) samples of the complexes were purified, as demonstrated by both SDS-PAGE and WB analyses (chromatograms, SDS-PAGE and WB analyses for complexes S, M, and L are shown in *Figures 9.6 - 9.9*) and by mass spectrometry (data not shown). The purification protocol of complex S included an additional purification stage of size exclusion chromatography (SEC) (the chromatogram

Sample	Studied protein-protein complexes	Sequences
Complex S	hYAP1 (50-120)	(-21)MGSSHHHHHHSQDPLVPR↓GSA(1) GHQIVHVRGDSETDLEALFNAVMNPKTANVPQTVPMRLRKLPDSFFKPPEPKSHSRQASTDAGTAGALTP(70)
	hTEAD4 (217-434)	(1)MRSVASSKLWMLEFSAFLEQQQDPDTYNKHLFVHIGQSSPSYSDPYLEAVDIRQIYDKFPEKKGGLKDLFER GPSNAFFLVKFWADLNTNIEDEGSSFYGVSSQYESPENMIITCSTKVCSFGKQVVEKVETEYARYENGHYSYRIH RSPLCEYMINFIHKLKHLPEKYMMNSVLENFTILQVVTNRDTQETLLCIAYVFEVSASEHGAQHHIYRLVKE(219)
Complex M	hYAP1 (50-140)	(-21)MGSSHHHHHHSQDPLVPR↓GSA(1) GHQIVHVRGDSETDLEALFNAVMNPKTANVPQTVPMRLRKLPDSFFKPPEPKSHSRQASTDAGTAGALTPQHVR AHSSPASLQLGAVSPG(90)
	hTEAD4 (217-434)	(1)MRSVASSKLWMLEFSAFLEQQQDPDTYNKHLFVHIGQSSPSYSDPYLEAVDIRQIYDKFPEKKGGLKDLFERG PSNAFFLVKFWADLNTNIEDEGSSFYGVSSQYESPENMIITCSTKVCSFGKQVVEKVETEYARYENGHYSYRIHR SPLCEYMINFIHKLKHLPEKYMMNSVLENFTILQVVTNRDTQETLLCIAYVFEVSASEHGAQHHIYRLVKE(219)
Complex L	hYAP1 (50-171)	(-21)MGSSHHHHHHSQDPLVPR↓GSA GHQIVHVRGDSETDLEALFNAVMNPKTANVPQTVPMRLRKLPDSFFKPPEPKSHSRQASTDAGTAGALTPQHVR AHSSPASLQLGAVSPGTLTPTGVVSGPAATPTAQHLRQSSFEIPDDV(121)
	hTEAD4 (217-434)	(1)MRSVASSKLWMLEFSAFLEQQQDPDTYNKHLFVHIGQSSPSYSDPYLEAVDIRQIYDKFPEKKGGLKDLFERG PSNAFFLVKFWADLNTNIEDEGSSFYGVSSQYESPENMIITCSTKVCSFGKQVVEKVETEYARYENGHYSYRIHR SPLCEYMINFIHKLKHLPEKYMMNSVLENFTILQVVTNRDTQETLLCIAYVFEVSASEHGAQHHIYRLVKE(219)

Table 9.2: Summary of the amino acid sequences of the S, M, and L complex investigated in this work. Amino acids belonging to the His^6-tag and to the thrombin recognition site are colored green and red, respectively, while the sequences of hTEAD4 YBD and hYAP1 TBD are in black.

and SDS-PAGE is displayed in *Figure 9.7*), showing that only the heterodimer is present in solution (retention volume compatible with the molecular weight of the complex S). Likewise, the results achieved by native mass-spectrometry and native PAGE analyses for complex L (*Figure 9.10*) revealed that in solution only the heterodimeric state is present (determined mass compatible with the molecular weight of the complex L).

The final purification yields obtained for the S, M, and L complexes were estimated to 60 mg L^{-1}, 30 mg L^{-1}, and 90 mg L^{-1} for bacterial culture, respectively.

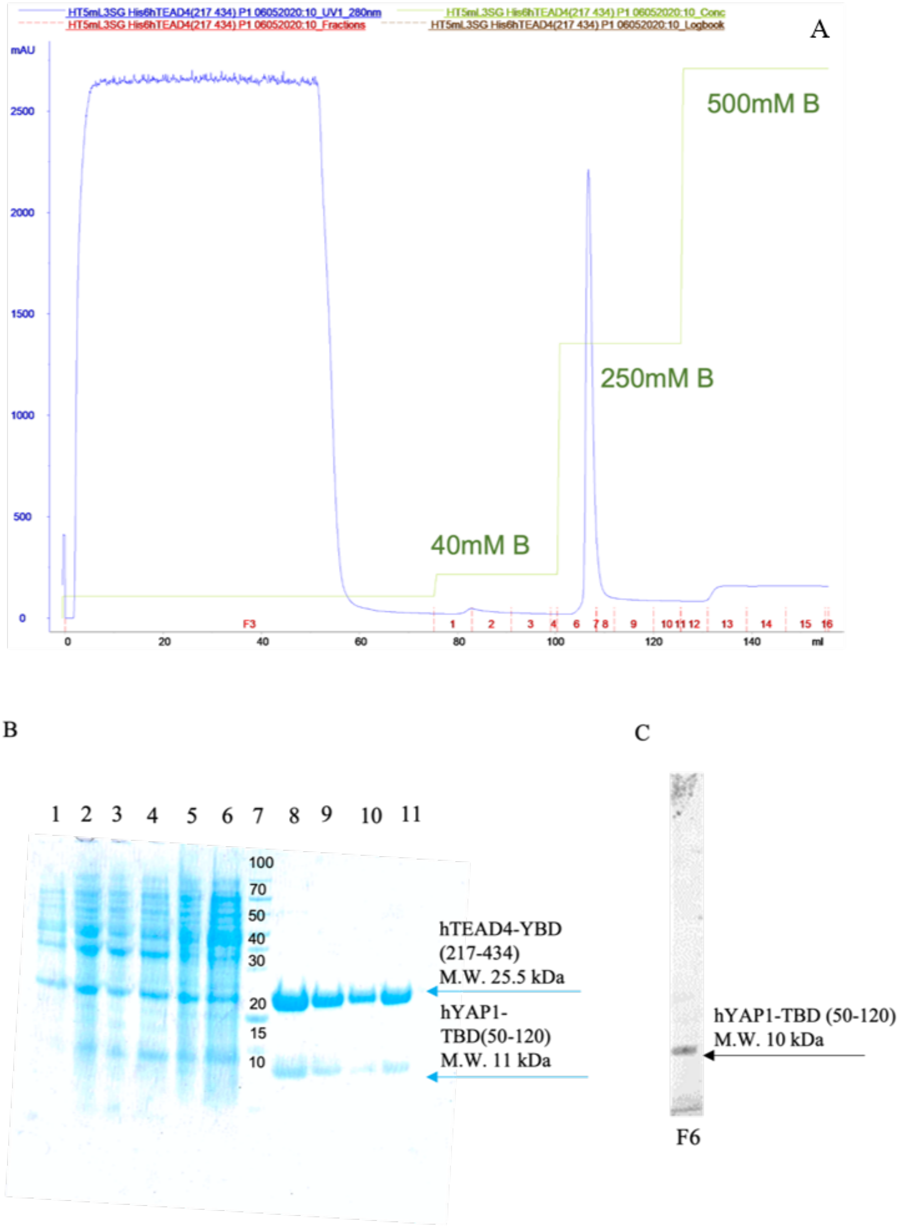


Figure 9.6: (**A**) Chromatogram of the nickel-affinity chromatography (HisTrap FF 5mL column, GE-Healthcare) purification of the S complex. The UV$_{280\text{nm}}$ monitoring and imidazole concentration are shown as blue and green lines, respectively. Elution fractions are indicated by red marks on the x-axis. (**B**) SDS-PAGE analysis of the fractions collected from the first stage of nickel-affinity chromatography (the protein ladder is in lane 7). (**C**) Western-blot analysis of the F6 fraction. His^6-tagged proteins are detected using (using the monoclonal anti-6x His-tag – HRP conjugated, Sigma Aldrich)

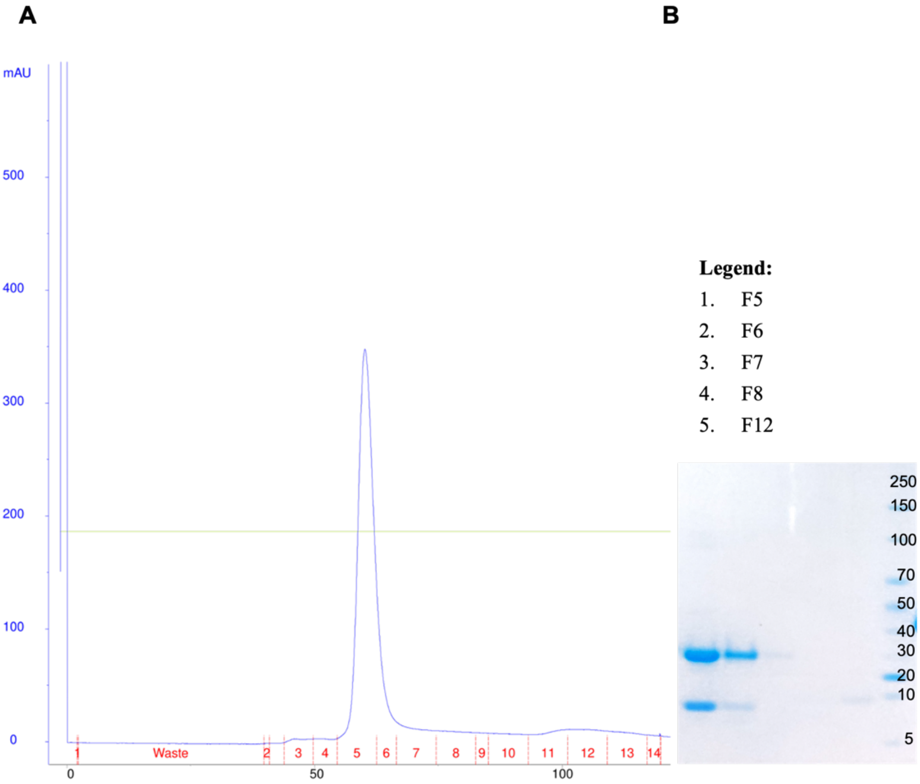


Figure 9.7: (A) Chromatogram of the HiLoad 16/600 Superdex 75pg column (GE-Healthcare) purification step. The UV$_{280\mathrm{nm}}$ monitoring of protein elution is in blue; fractions are indicated by red marks on the x-axis. (**B**) SDS-PAGE analysis of the fractions collected from SEC. The protein ladders are in lane 6.

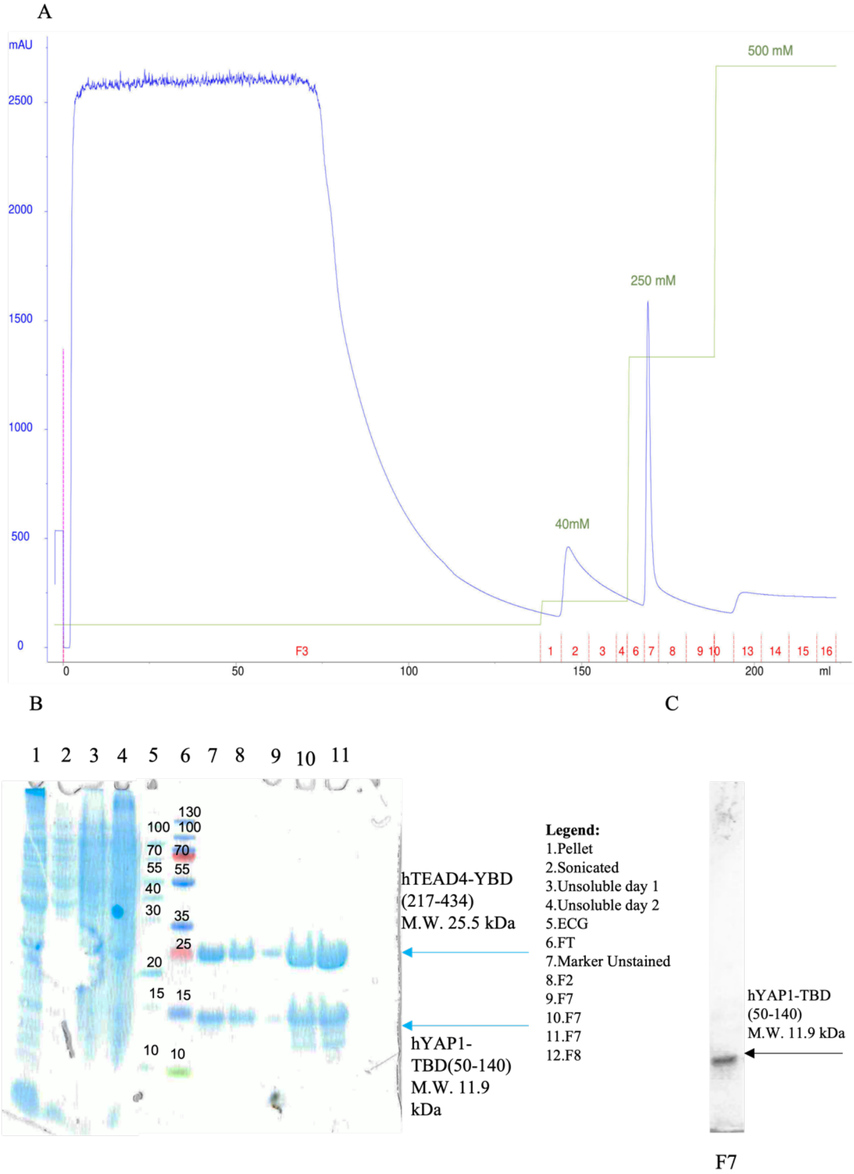


Figure 9.8: (**A**) Chromatogram of the complex M purification through nickel affinity chromatography (HisTrap FF 5mL column, GE-Healthcare). The UV$_{280\text{nm}}$ monitoring and imidazole concentration are shown as blue and green lines, respectively, whereas; the elution fractions are indicated by red marks on the x-axis. (**B**) SDS-PAGE analysis of the fractions collected from the first stage of Ni^{2+} affinity chromatography. The protein ladder is in lane 7. (**C**) Western-blot analysis of the F7 fraction. His^{6}-tagged proteins are detected using the (using the monoclonal anti-6x His-tag – HRP conjugated, Sigma Aldrich)

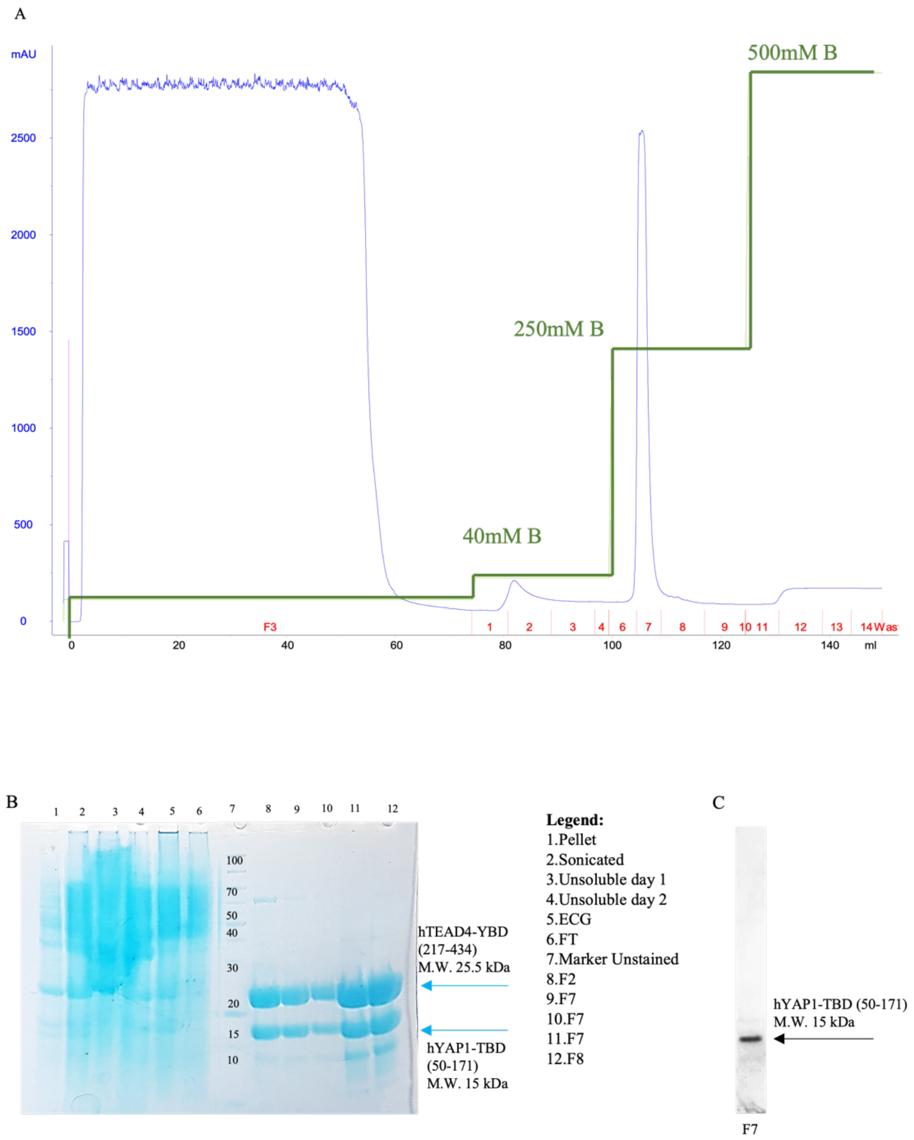


Figure 9.9: (**A**) Chromatogram of the complex L purification through nickel affinity chromatography (HisTrap FF 5mL column, GE-Healthcare). The UV_{280nm} monitoring and imidazole concentration are shown as blue and green lines, respectively, whereas; the elution fractions are indicated by red marks on the x-axis. (**B**) SDS-PAGE analysis of the fractions collected from the first stage of Ni^{2+} affinity chromatography. The protein ladder is in lane 7. (**C**) Western-blot analysis of the F7 fraction. His^{6}-tagged proteins are detected using an HRP-conjugated antibody able to recognize the His6 epitope.

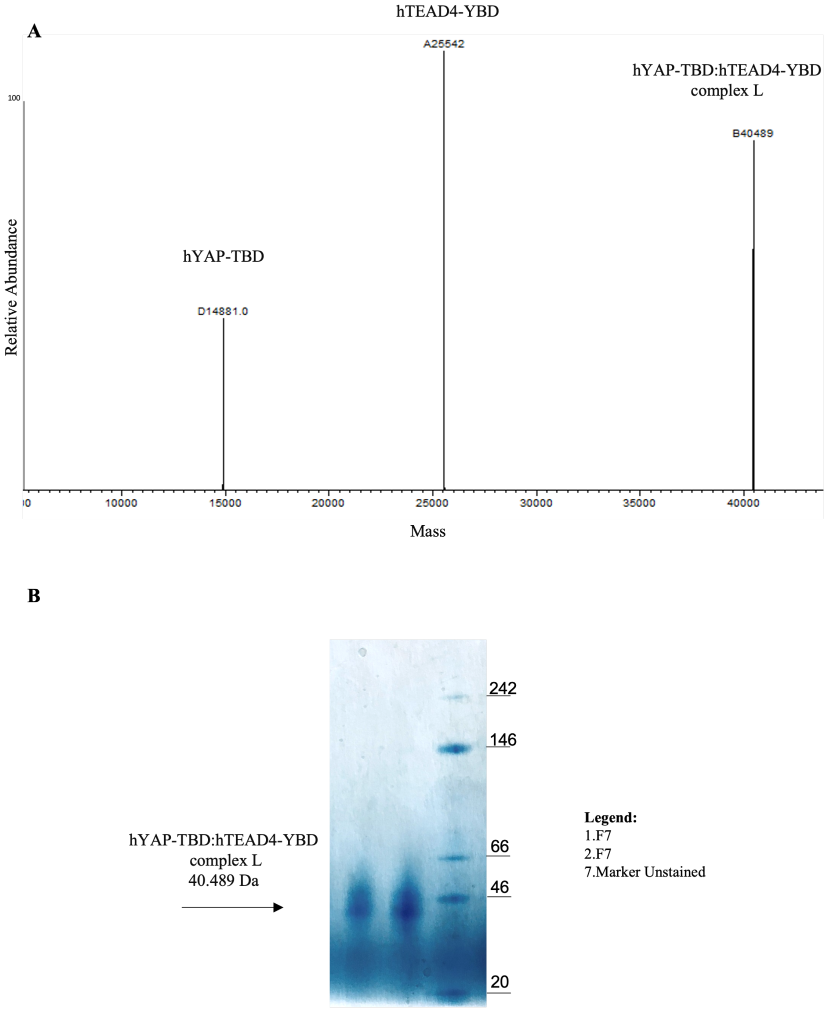
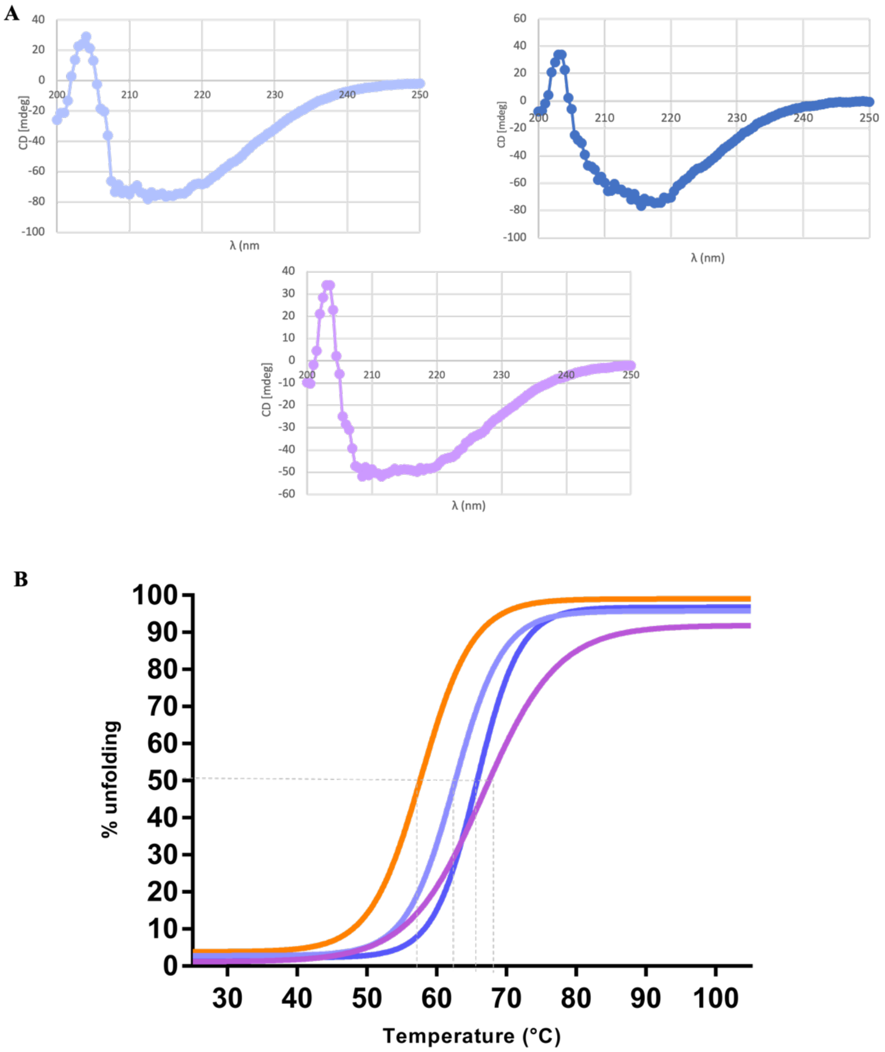


Figure 9.10: (**A**) Mass spectrum in native conditions of fraction F7 of complex L, His6 – hYAP TBD:hTEAD4 YBD, purified by nickel- affinity chromatography, (hYAP TBD S theoretical MW = 15014Da, determined MW = 14881 Da, hTEAD4 YBD theoretical MW = 25543 Da, determined MW = 25542 Da, hYAP TBD:hTEAD4 YBD S complex theoretical MW = 40540 Da, determined MW = 40489 Da).(**B**) Native-PAGE analysis of fractions F7 and F5 of complex L, His6 – hYAP TBD:hTEAD4 YBD, separated by nickel-affinity chromatography. The migration of the single band of fraction F7 corresponds to a MW of 40.489 Da, indicating the presence of a heterotetrametric assembly in solution.

9.2.2 Thermal stability analysis of the S, M, and L complexes

The CD spectra recorded on complexes S, M and L are consistent with folded proteins (*Figure 9.11 A*). The formation of a protein-protein complex can affect the stability of each partner proteins; the assessment of such stabilization effects can be performed by analyzing the thermal stability of the complexes. On this purpose, CD thermal denaturation studies have been carried out on the S, M and L complexes, by monitoring the CD signal at 220 nm in the temperature range 25-105 °C. Unfolding transition curves were thus calculated (*Figure 9.11 B*) and analyzed to determine T_m values. The lowest the T_m value, 62.55 (±0.1) °C, was obtained for the S complex (light purple unfolding transition curve in *Figure 9.11 B*), whereas values of 65.73 (±0.1) °C and 66.68 (±0.2) °C, were calculated for the M and L complexes, respectively (dark violet and lilac unfolding transition curves in *Figure 9.11 B*). Thus, the interaction of hTEAD4 YBD with the hYAP TBD S fragment (50-120) results in a T_m increase of ≈4.0 °C with respect to the isolated hTEAD4 YBD (*section 9.1.2*). Our finding agrees with the data of *Mesrouze et al*, which reported shifts of ≈5.0 °C and ≈4.2 °C in the T_m of hTEAD4 YBD and Myr-hTEAD4 YBD, respectively, in presence of syn-hYAP TBD 60-99. Moreover, increasing the length of the hYAP TBD from the S to the L fragments, leads to complexes with improved thermal stability, as demonstrated by the increment of ≈4.0 °C in their T_m values (*Figure 9.11 B*). These data seem to suggest that more extended areas of the hYAP TBD, beyond the region 60-99 of the synthetic peptide, are involved in the interaction with hTEAD4 YBD. Notably, the highest thermal stability is observed for the L complex, in which the whole hYAP TBD, fragment 50-171, is employed. Indeed, the fragment 50-171 was mapped biochemically by *Vassilev et al* [154] as the one accounting for the interaction with hTEAD4.



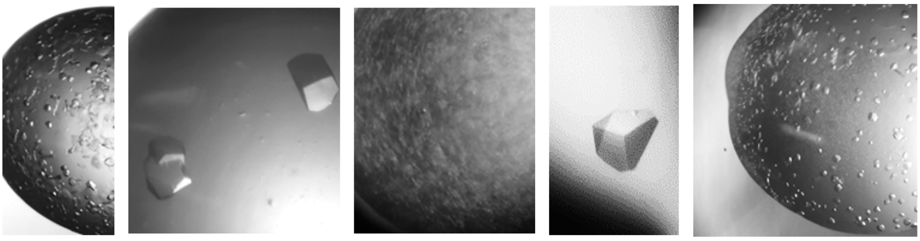
Protein sample	hTEAD4-YBD	hYAP-hTEAD4 S complex	hYAP-hTEAD4 M complex	hYAP-hTEAD4 L complex
Melting Temperature (Tm) °C	57.82 ± 0.1	62.55 ± 0.1	65.73 ± 0.1	66.68 ± 0.2

Figure 9.11: In the upper panel (**A**) CD curves for S, M and L complexes. In the lower panel (**B**), thermal unfolding transition curves determined for the S, M and L complexes by CD thermal denaturation analysis. Measurements were performed in triplicate. Melting temperature (T_m) values determined for each complex are reported in the table.

9.2.3 Crystallization of hYAP TBD:hTEAD4 YBD S, M, and L complexes

The huge crystallization screening campaign performed on the His^6-hYAP TBD:hTEAD4 YBD S, M and L complexes, allowed us to observe crystal growth in various conditions, showing their propensity to crystallize in presence of PEG 3350 (*Table 9.12, 9.13 and 9.14*). Although in different conditions, this agent was also employed in the crystallization of the syn-hYAP TBD 60-99:hTEAD4 YBD complex [156, 161, 187]. All successful conditions were further optimized by varying protein, PEG and salt concentrations, buffer pH, and incubation temperature, nonetheless, preliminary X-ray crystallographic analyses revealed only poor diffraction patterns for these crystals, with resolutions >10 Å. Thus, we decided to move to the micro-seeding crystallization technique, aiming to improve the crystal ordering and diffraction quality. The most prominent improvement in the crystal properties was achieved for the S complex for which we obtained data to 2.5 Å. These crystals grew in 48 hours, displaying a hexagonal bipyramidal habit (*Figure 9.15*). On the other hand, maximum resolution of only 6 Å and 3.1 Å were obtained, for the M and L complexes, despite the application of the micro-seeding technique also to grow these crystals.

hYAP TBD:hTEAD4 YBD complex S



0.2 M sodium malonate pH 5.0 + 15% w/v PEG 3350

4% tacsimate pH 5.0 + 12% w/v PEG 3350

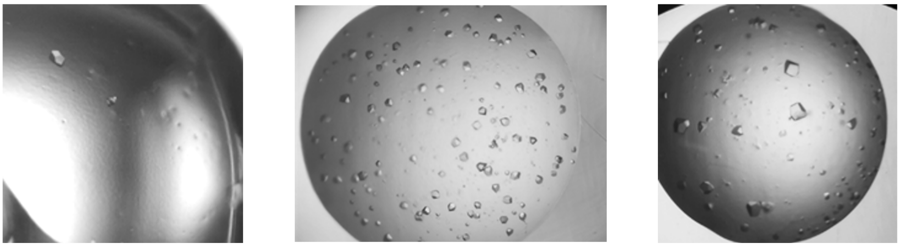
8% tacsimate pH 5.0 + 20% PEG 3350

0.1 M sodium malonate pH 5.0 + 12% w/v PEG 3350

0.2 M calcium chloride dihydrate + 20% w/v PEG 3350

Figure 9.12: Pictures and precipitant composition of crystals of complex

hYAP TBD:hTEAD4 YBD complex M



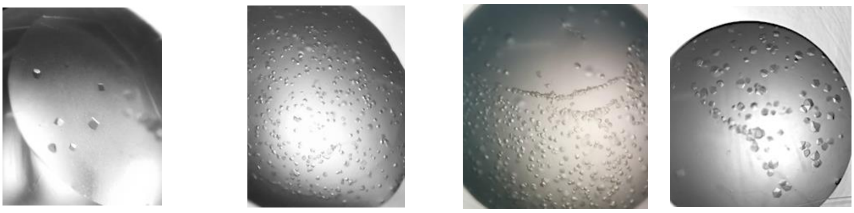
0.2 M sodium malonate pH 5.0, 12% w/v PEG 3350

0.2 M ammonium chloride, 20% w/v PEG 3350

0.2 M ammonium chloride, 20% w/v PEG 3350

Figure 9.13: Pictures and precipitant composition of crystals of complex M

hYAP TBD:hTEAD4 YBD complex L

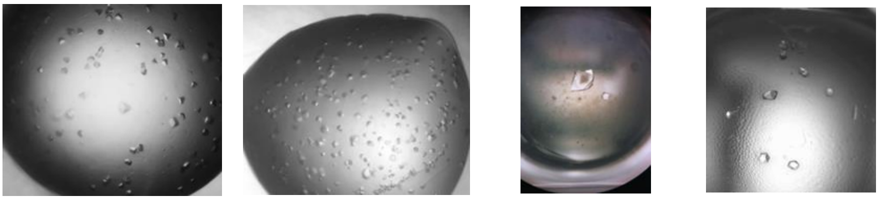


0.2 M calcium chloride dihydrate, 20% w/v PEG 3350

0.2 M sodium chloride, 20% w/v PEG 3350

0.2 M potassium chloride, 20% w/v PEG 3350

0.2 M ammonium chloride, 20% w/v PEG 3350

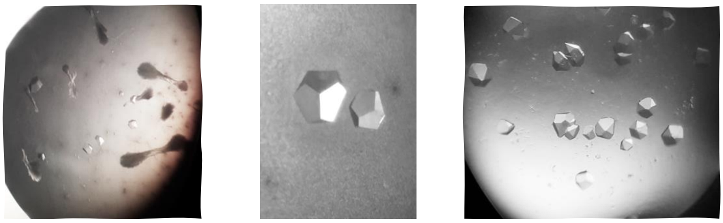


0.2 M potassium thiocyanate, 20% w/v PEG 3350

0.2 M lithium nitrate, 20% w/v PEG 3350

0.2 M hydrated calcium acetate, 20% w/v PEG 3350

0.2 M sodium malonate pH 6.0 + 20% w/v PEG 3350



0.2 M sodium malonate pH5+20% PEG 3350

0.2 M sodium malonate pH5+20% PEG 3350+ seeding solution 1:1000

Figure 9.14: Pictures and precipitant composition of crystals of complex L

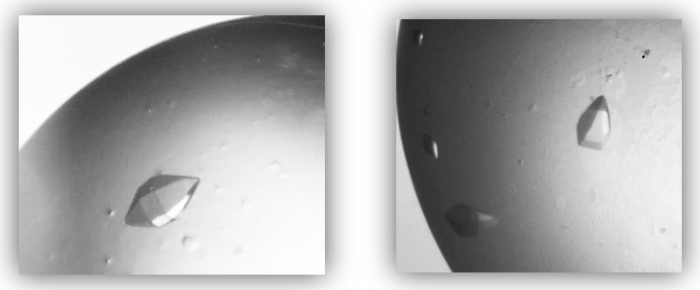


Figure 9.15: Crystals of complex S grown using a precipitant solution composed of 0.2 M calcium chloride dihydrate and 15 % PEG 3350 and seeding solution having a 1:1000 dilution factor.

9.2.4 Overall structure of YAP TBD:TEAD YBD S complex

The structure of the hYAP TBD:hTEAD4 YBD S complex has been determined to 2.5 Å resolution, in the trigonal space group $P3_121$ with unit cell parameters a,b = 164.58 Å, c = 258.61 Å. The crystal ASU consists of 12 hYAP TBD S fragment: hTEAD4 YBD heterodimers (*Figure 9.16*). In all heterodimers, the hTEAD4 YBD chain was completely traced apart for the segment including residues 254-261. On the other hand, the electron density observed for the hYAP1 TBD S fragment allowed us to rebuild only the segment including residues 51-100. The starting N-terminal amino acid (residue 50) and the C-terminal segment 101-120 of the hYAP1 TBD S fragment, and the additional twelve amino acids belonging to the N-terminal His^6-tag were not traced in our model, reasonably because of their positional disorder. The hTEAD4 YBD adopts an immunoglobulin-like fold, with a mix of α/β fold (*Figure 9.17*). The 12 β-strands forming two β-sheets packed against each other, whereas the 4 α-helices are organized around them. In detail, the first β-sheet is constituted by strands $\beta1$, $\beta2$, $\beta5$, $\beta8$, and $\beta9$; while the second is formed by strands $\beta3$, $\beta4$, $\beta6$, $\beta7$, $\beta10$, $\beta11$, and $\beta12$. hTEAD4 YBD shows two helix-turn-helix motifs, the first involves $\alpha1$ and $\alpha2$, connecting strands $\beta3$ and $\beta4$ and the second consisting of $\alpha3$ and $\alpha4$ which link strands $\beta9$ and $\beta10$. The analysis of each hTEAD4 YBD chain reveals the presence of a myristic acid molecule populating their hydrophobic pocket (detailed in *section 9.2.6*). The hYAP TBD S fragment wraps around the hTEAD4 YBD (*Figure 9.17*). *Li et al* [161] reported that the isolated hYAP TBD is largely unstructured and it folds only in presence of hTEADs by wrapping around it. In the structure of the S complex, three main secondary structural elements characterize the hYAP TBD S fragment: the N-terminal β-strand $\beta1$, followed by the α-helix $\alpha1$, and the Ω-loop (Figure 53). All these elements take direct interactions with hTEAD4 YBD (as detailed in the following section). The α-helix $\alpha1$ and the Ω-loop are connected by a 10-amino acid flexible linker region including the PXXΦP motif (P: proline; X: any residue; Φ: hydrophobic residue), shown to be important for the interaction with hTEADs [44, 163]. Indeed, the flexibility of this linker seems to allow the optimal placement of the hYAP $\alpha1$ and Ω-loop, forming the two main interface areas connecting the partner proteins [44, 161].

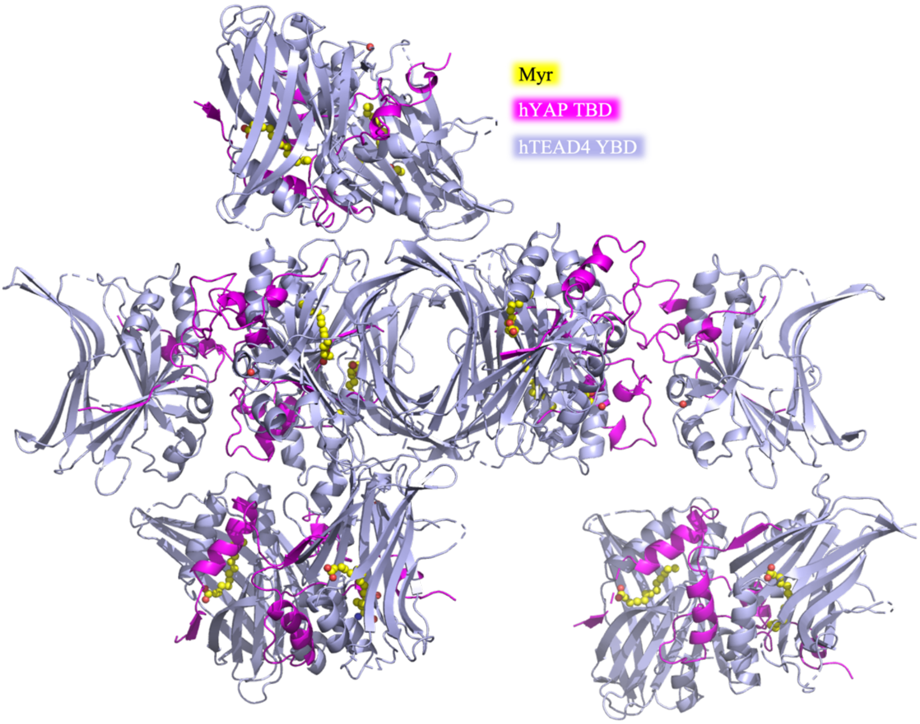


Figure 9.16: Cartoon representation of the overall structure of complex S The 12 hYAP TBD:hTEAD4 YBD complexes are shown. hYAP TBD and hTEAD4 YBD are represented as light-blue and magenta cartoon, respectively. The myristic acid molecules, populating the hydrophobic pockets of all hTEAD4 YBD, are displayed in spere, yellow carbon atoms.

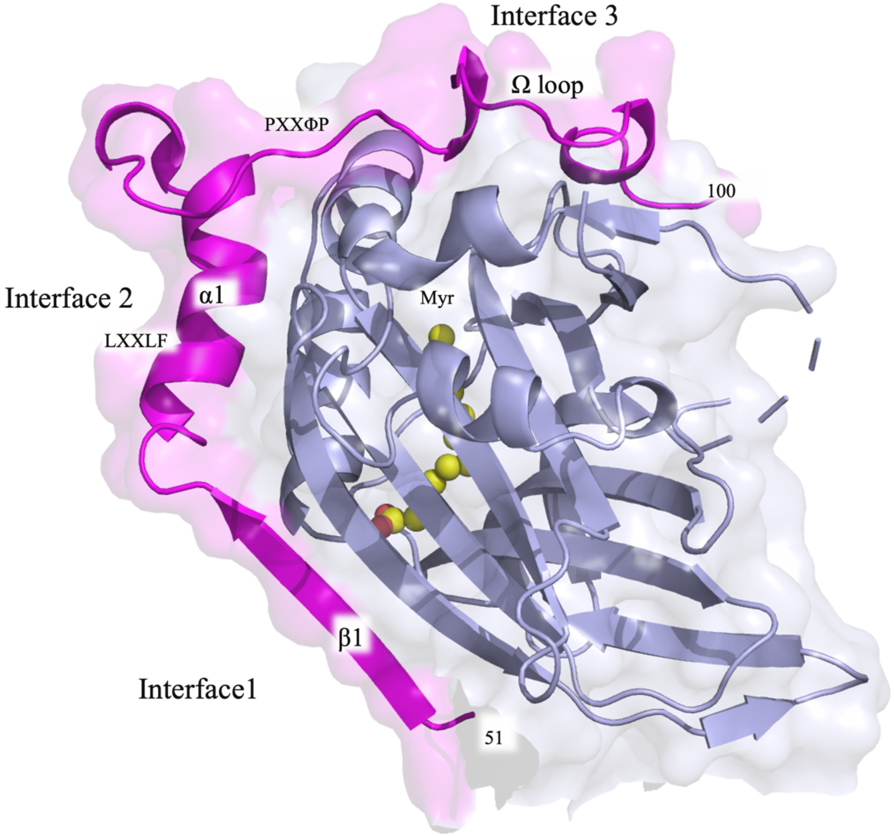


Figure 9.17: Cartoon representation of a representative heterodimer of hYAP TBD:hTEAD4 YBD, complex S. hYAP TBD and hTEAD4 YBD are represented light-blue and magenta cartoon, respectively. The motifs LXXLF and PXXΦP, and the three interaction regions $\alpha1$, $\beta1$, and Ω loop of hYAP TBD are labelled in the Figure. The myristic acid molecule, populating the hydrophobic pocket of hTEAD4 YBD, is displayed as sphere, yellow carbon atoms.

9.2.5 Analysis of the main PPIs characterizing the hYAP TBD:hTEAD4 YBD S complex

The structure of the S complex shows that the hYAP TBD S fragment interacts with hTEAD4 YBD via the formation of three interface regions. The first interface is formed between the strands $\beta1$ of hYAP TBD and $\beta7$ of hTEAD4

YBD, extending the β-sheet of the latter protein by one antiparallel strand formed by the partner protein (*Figure 9.18*). A network of five intermolecular H-bonds is established between the peptide backbones of residues Ile54, His56, and Arg58 on the hYAP TBD $\beta1$, and Val 341, Glu343, and Val345 of the hTEAD4 YBD $\beta7$. Moreover, the hTEAD4 YBD Thr347 takes a water-mediated interaction with the hYAP TBD His52 (*Figure 9.18*). At variance with interfaces 2 and 3, interface 1 is mainly formed between backbone regions of the partner proteins, limiting the contribution of the amino acid composition of this hYAP area, which is less conserved across species [161]. The second interface is formed between the hYAP TBD 1 that packs into the shallow groove lined by hTEAD4 YBD $\alpha3$ and $\alpha4$ (*Figure 9.19*). Several hydrophobic interactions occur between the hYAP TBD residues Leu65, Leu68, and Phe69 and the hTEAD4 YBD residues Phe337, Tyr369, Phe373, Lys376, Leu377, Leu380, Val389, Asn392 and Phe393, making this a hydrophobic interface area (*Figure 9.19*). Here, the hYAP TBD shows the conserved LXXLF motif ($L^{65}XXL^{68}F^{69}$), being a peculiar binding motif to hydrophobic grooves [188, 189]. The third interface involves the hYAP TBD Ω-loop and the hTEAD4 YBD hydrophobic pocket lined by residues belonging to its $\alpha1$, $\alpha4$, $\beta11$, and $\beta12$ (*Figure 9.19*). Residues Met86, Leu91, and Phe95 of the hYAP TBD are implicated in van der Waals interactions with hTEAD4 YBD Val265, Leu295, Lys297, Val414 and Tyr429 (*Figure 9.19*). Furthermore, these hYAP TBD residues interact also with each other, forming the hydrophobic core of the Ω-loop. The following residue on the Ω-loop, Phe96, forms additional intramolecular interactions with Arg87, Phe95 and Leu91 contributing to stabilize the folding of this hYAP TBD area. Interface 3 is also strengthened by hydrophilic intermolecular interactions entailed between the partner proteins. The guanidinium group of hYAP TBD Arg89 is H-bonded to the carboxylate of hTEAD4 YBD Asp272. Furthermore, the hydroxyl of hYAP TBD Ser94 is H-bonded to hTEAD4 YBD Glu263 and Tyr429 (*Figure 9.20*). Finally, the backbone nitrogen of hYAP TBD Met86 forms H-bond with the carboxylate of hTEAD4 YBD Glu391 (*Figure 9.20*). Interface 3 has a pivotal role for the YAP:TEAD interaction, indeed the mutation of the residues 94 on the hYAP

Ω-loop, as well as those of Glu263 and Tyr429 on hTEAD4 YBD has a drastic effect on the complex formation [163]. Up to now, the structural information on the hYAP:hTEAD4 complex was limited to the complex of hTEAD4 YBD with syn-hYAP TBD 60-99 (PDB id 5OAQ [156]). This complex shows only two contact areas between the partner proteins, namely interfaces 2 and 3. The syn-hYAP TBD 60-99 does not include the N-terminal segment required to form the $\beta 1$ structural element on hYAP, preventing the formation of this interacting interface. On the contrary, this contact is observed in our complex that thus provides a more comprehensive information on the hYAP:hTEAD4 PPI. The comparison with the syn-hYAP TBD 60-99:hTEAD4 YBD complex shows that, despite the difference observed at the interface 1, the other two contact areas, interfaces 2 and 3, are conserved in both structures (*Figure 9.21*). Although never reported in the previous crystal structures of the hYAP TBD:hTEAD4 YBD complexes (PDB id 5OQA [156] and PDB id 6GE3), the formation of interface 1 was formerly observed in the structure of complex of hTEAD1 YBD with hYAP1 (residues 51-100) (PDB id 3KYS) by *Li et coworkers* [161]. The comparison with our structure of the S complex shows a highly retained binding mode of the hYAP TBD around the hTEADs YBD, and the formation of the same three interacting interfaces (*Figure 9.22*). A deeper analysis of these contact areas in both complexes evidences those analogous interactions are formed between the partner proteins in spite of the hTEAD isoform. This is consistent with the high sequence homology of the hTEADs YBD, resulting in an identity of 76.5%, (*Figure 7.3*) [148]. Notably, the hTEAD YBD residues involved in the interaction with hYAP TBD are highly conserved between isoforms 1 and 4 [161, 190], supporting their prominent role in these PPIs.

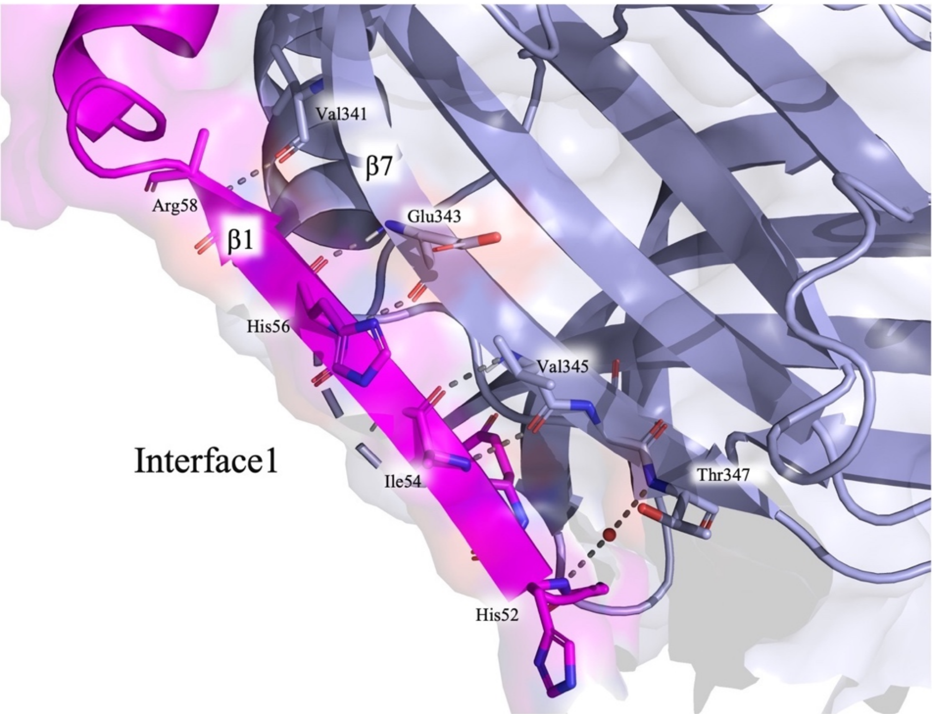


Figure 9.18: Representation of interface 1 between hYAP TBD (magenta cartoon):hTEAD4 YBD (light-blue cartoon). H-bonds (grey dashed lines) between the peptide backbones of hYAP TBD $\beta1$ (residues 54-57, (in stick, magenta carbon atoms) and hTEAD4 YBD $\beta7$ (residues 341-345 in stick, light-blue carbon atoms) are shown. Water-mediated interaction (grey dashed lines) between hTEAD4 YBD Thr347 and hYAP TBD His52 in the interface 1. The water molecule is represented as a red sphere.

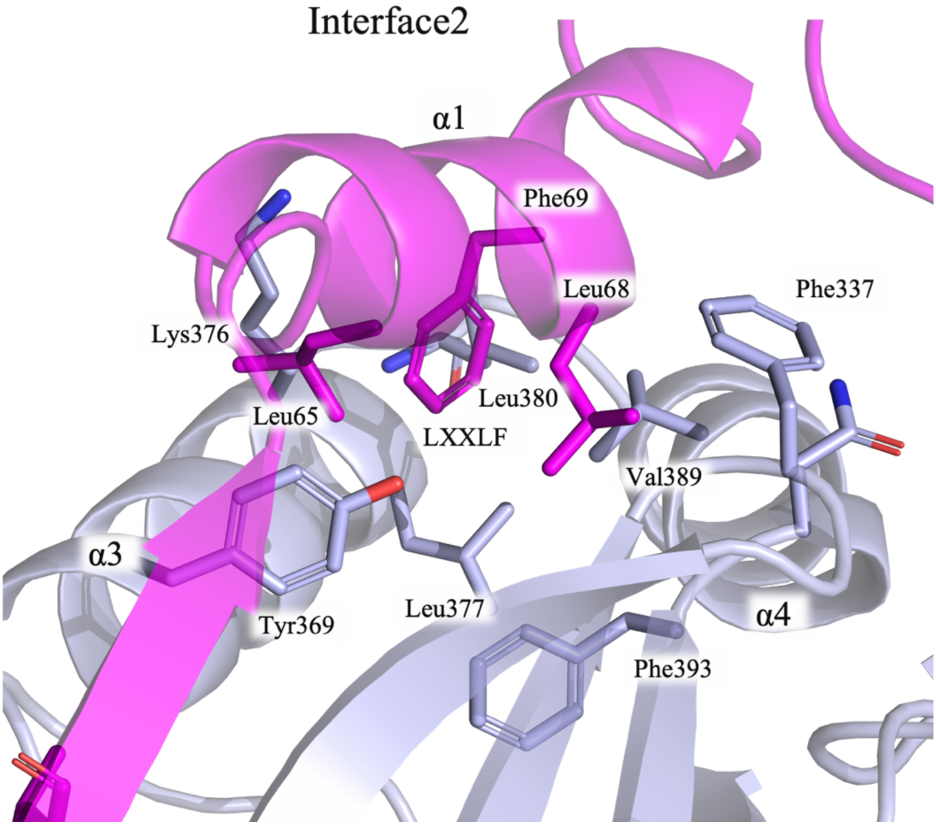


Figure 9.19: Second interface between hYAP TBD (magenta cartoon and carbon atoms) and hTEAD4 YBD (light-blue cartoon and carbon atoms). The hYAP TBD $\alpha 1$ helix protrudes into the hTEAD4 YBD binding groove formed by $\alpha 3$ and $\alpha 4$, forming several hydrophobic interactions. The hYAP TBD residues Leu65, Leu68, and Phe69 ($L^{65}XXL^{68}F^{69}$ motif) and the hTEAD4 YBD residues Phe337, Tyr369, Phe373, Lys376, Leu377, Leu380, Val389, and Phe393 are involved in the formation of interfaces 2 and they are shown in the Figure.

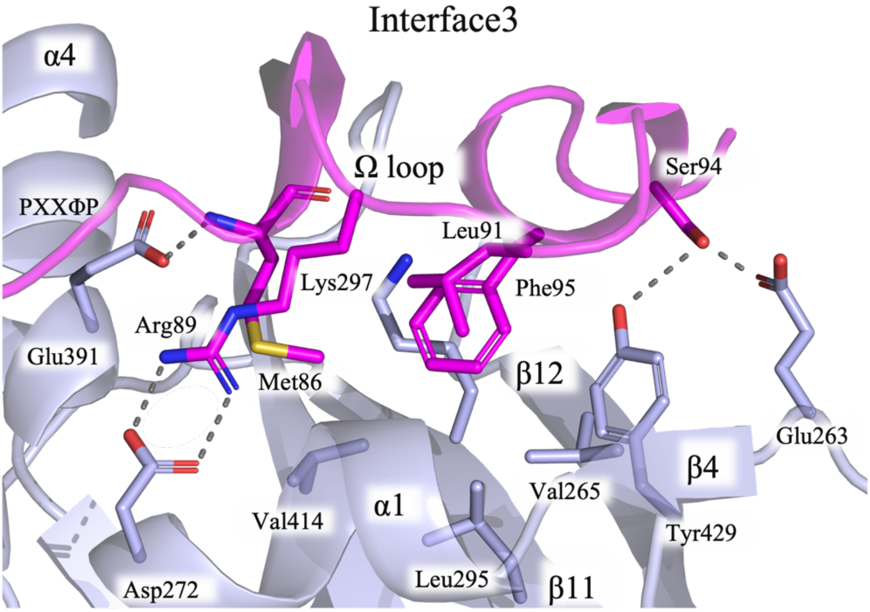


Figure 9.20: Third interface between hYAP TBD (magenta cartoon and carbon atoms) and hTEAD4 YBD (light-blue cartoon and carbon atoms). This interface involves the hYAP TBD Ω-loop and the hTEAD4 YBD hydrophobic pocket lined by residues belonging to its $\alpha 1$, $\alpha 4$, $\beta 11$, and $\beta 12$. Several van der Waals interactions occur between the Met86, Leu91, and Phe95 of the hYAP TBD and Val265, Leu295, Lys297, Val414 and Tyr429 of hTEAD4 YBD (the mentioned residues are shown in the Figure). In addition, some H-bonds (grey dashed lines) are established between hYAP TBD Arg89, Met 86, Ser 94 (in stick, magenta carbon atoms) and hTEAD4 YBD Asp272, Glu391, Tyr429, Glu263 (in stick, light-blue carbon atoms)

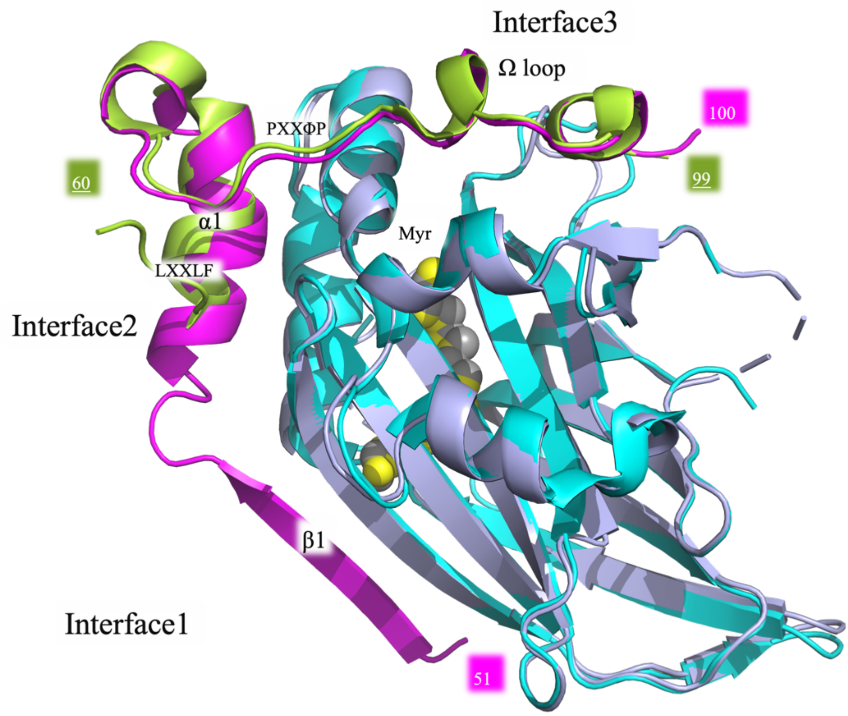


Figure 9.21: Cartoon representation of the superimposition between hYAP TBD:hTEAD4 YBD heterodimer complex S (hYAP TBD in magenta, hTEAD4 YBD in light-blue) and syn-hYAP TBD 60-99:hTEAD4 YBD (hYAP TBD in lemon-green and hTEAD4 YBD in cyan) (PDB id 5OAQ) [156]. The three interaction regions $\alpha 1$, $\beta 1$, and Ω loop of hYAP TBD are labelled in the Figure In this representation is evident the presence of an extra interaction interface, interface 1, in our hYAP TBD:hTEAD4 YBD complex S. The N and C-terminal residue numbers of hYAP TBD in our complex S (51-100) and of syn-hYAP TBD (60-99) are reported. Myristic acid in the structure of our hYAP TBD:hTEAD4 YBD complex S and that of hYAP TBD:hTEAD1 YBD are in sphere, yellow and grey respectively.

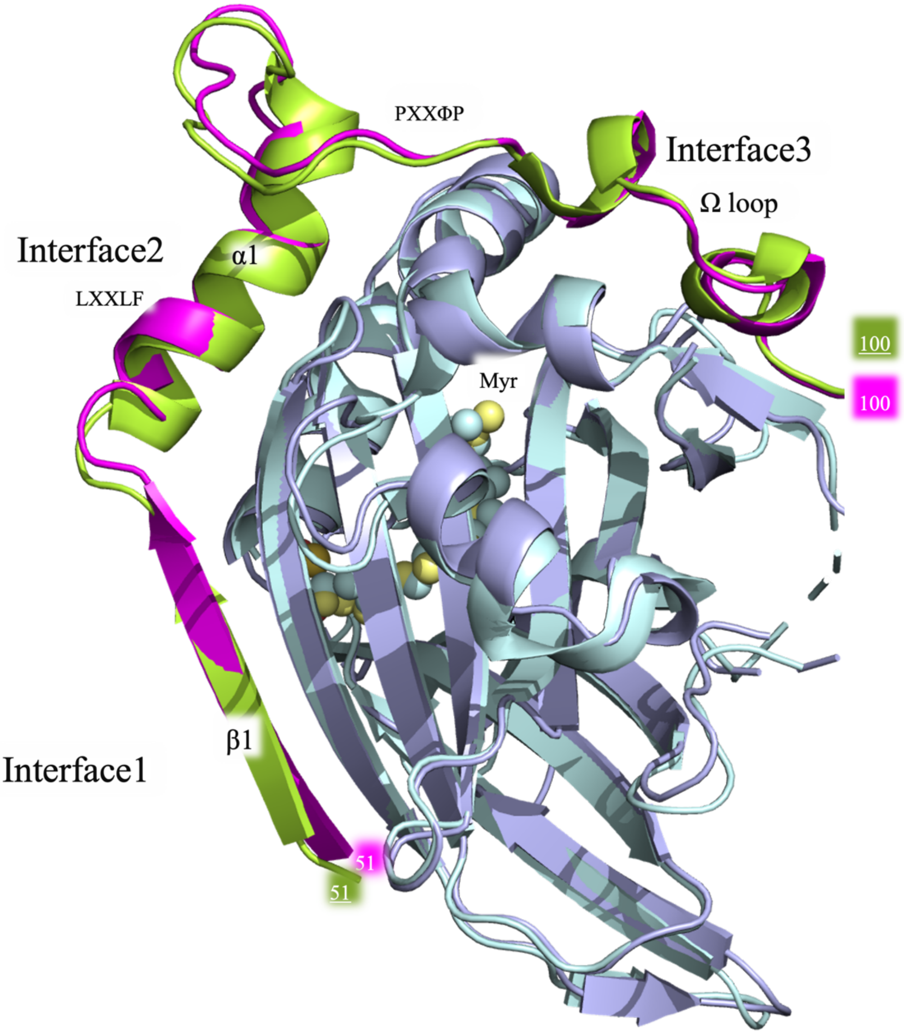


Figure 9.22: Cartoon representation of the superimposition between hYAP TBD:hTEAD4 YBD heterodimer complex S (hYAP TBD in magenta, hTEAD4 YBD in light-blue) and hYAP TBD:hTEAD1 YBD (hYAP TBD in lemon-green and hTEAD4 YBD in cyan) (PDB id 3KYS [156]). The three interaction regions $\alpha1$, $\beta1$, and Ω loop of hYAP TBD and the three interfaces between hYAP TBD and hTEAD4 YBD are labelled in the Figure. The N and C-terminal residue numbers of hYAP TBD in our compex S (51-100) and of syn-hYAP TBD (51-100) are reported. Myristic acid in the structure of our hYAP TBD:hTEAD4 YBD complex S and that of hYAP TBD:hTEAD1 YBD are in sphere, yellow carbon atoms and and grey carbon atoms, respectively

9.2.6 Myristic acid binding to hTEAD4 YBD in the structure of the S complex

In all heterodimers, a myristic acid was found to populate the hydrophobic pocket of hTEAD4 YBD. Since myristic acid was not added to the purified protein, its binding reasonably occurred within bacterial cells, during the production of the target proteins (myristic acid is produced under physiological conditions by both prokaryotic [191] and eukaryotic cells [192]). The molecule of myristic acid was modelled in each hTEAD4 YBD hydrophobic cavity, according to the observed electron density. The lack of a continuous electron density extending from the thiol of hTEAD4 YBD Cys367 suggests that myristic acid is not covalently bound to the protein (*Figure 9.23*) [192, 157]. The carboxylate head of myristic acid is instead H-bonded to the thiol and backbone nitrogen of Cys367 (*Figure 9.24 A*). The hydrophobic tail forms a network of van der Waals interactions with some hydrophobic residues exposed inside the pocket, mainly, Phe229, Val316, Thr332, Val334, Leu366, Met370, Leu390, Ile395 and Phe415 (*Figure 9.24 B*). The presence of myristic acid bound to hTEAD4 YBD was formerly reported also in the structure of the complex syn-hYAP TBD 60-99 (PDB ids 5OAQ [156] and 6GE3 [163]), where it acylated Cys367. However, despite the lack of a covalent linkage in our structure the binding mode of the fatty acid molecules inside the hTEAD4 YBD hydrophobic pocket is highly retained in all complexes (*Figure 9.25*). Thus, the occurrence of protein acylation at Cys367 seems not required for the population of this protein cavity by myristic acid, that can be driven by the hydrophobic nature of the pocket.

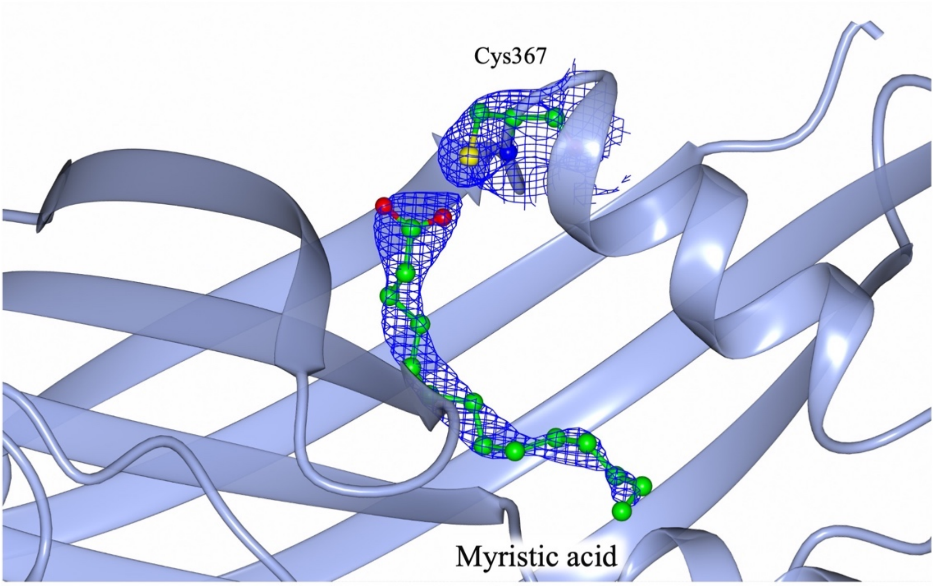


Figure 9.23: Hydrophobic central pocket view of hTEAD4 YBD (light-blue cartoon) populated by myristic acid (in sticks, carbon atoms are colored green) on the structure of our complex S. The Cys367 (in sticks, carbon atoms in green) is shown.

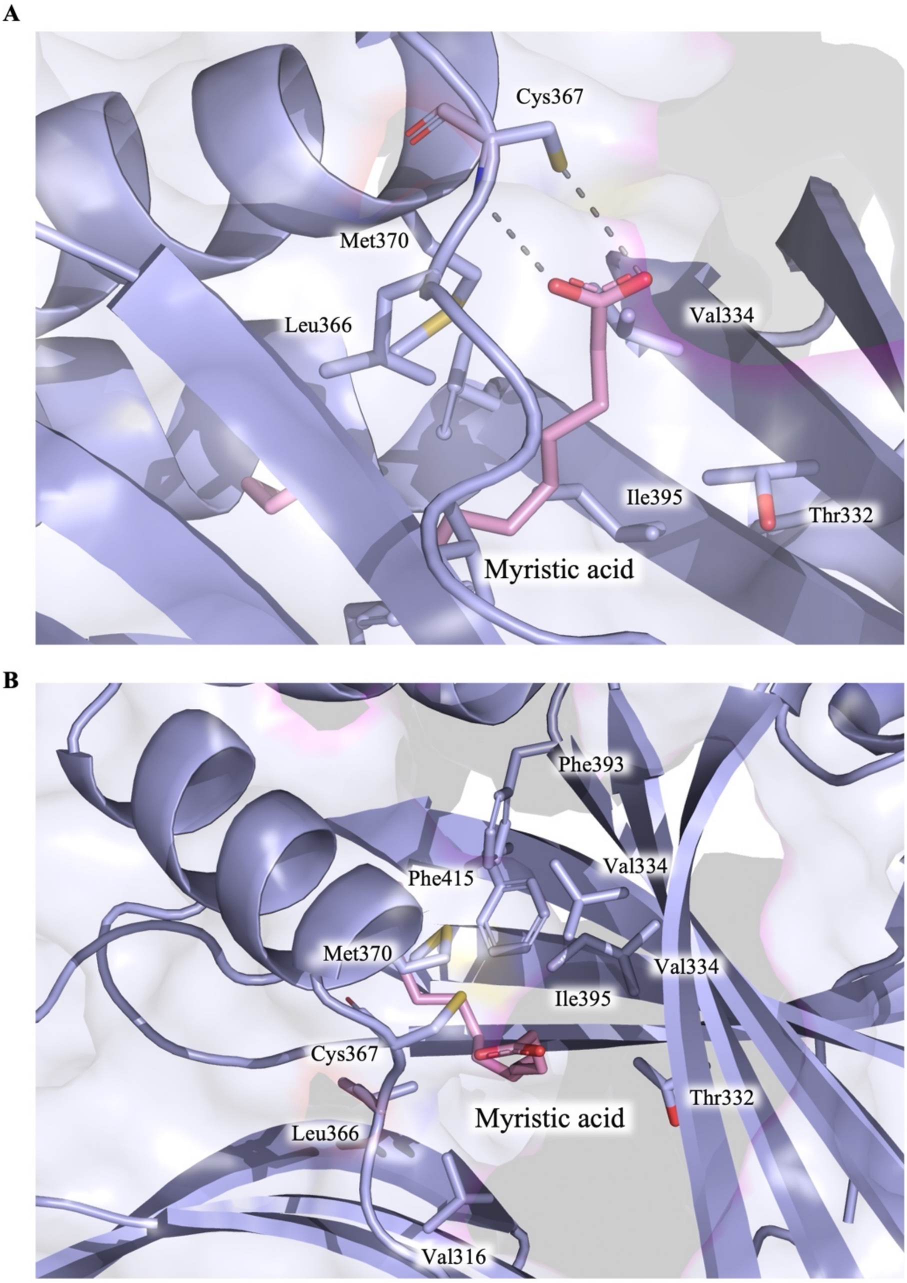


Figure 9.24: Hydrophobic central pocket view of hTEAD4 YBD (light-blue cartoon) populated by myristic acid (in sticks, carbon atoms are colored pink). (**A**) View of myristic acid H-bonded (grey dashed lines) to the thiol and backbone nitrogen of Cys367 (**B**) View of hydrophobic interactions occurring between the aliphatic chain of myristic acid (pink sticks) and the hTEAD4 YBD residues (Phe229, Val316, Thr332, Val334, Leu366, Met370, Leu390, Ile395 and Phe415) exposed inside the hydrophobic pocket (light-blue cartoon and carbon atoms, residues in sticks).

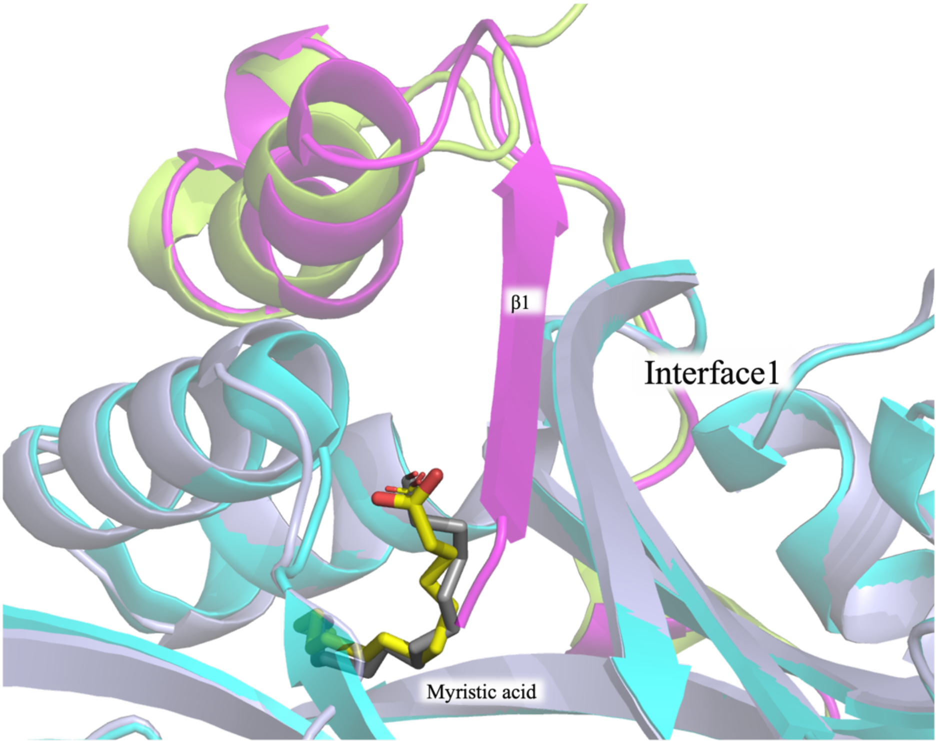


Figure 9.25: Cartoon representation of the superimposition between myristic acid binding pocket of hYAP TBD:hTEAD4 YBD heterodimer complex S (hYAP TBD in light blue, hTEAD4 YBD in magenta and myristic acid in sticks, pink carbon atoms) and myristoylated hTEAD4 in the syn-hYAP TBD 60-99 :hTEAD4 YBD complex (PDB id 5OAQ [156]) (syn-hYAP TBD 60-99 in light-pink, hTEAD4 YBD in violet and myristic acid in dark grey). The myristic acid populates the central hydrophobic pocket of hTEAD4 YBD, but in our complex S hTEAD4 YBD is not acylated.

9.3 Studies on full length hYAP and its complex with hTEAD4 YBD (XL complex)

9.3.1 Cloning and expression trials of full length hYAP

In the present study, the hYAP gene was cloned in two expression vectors, pMAL-p5X and pQE80L (*Figure 8.4*), allowing the expression of the target both as MBP-fusion protein and as His^6-tag hYAP (amino acid sequences are shown in *Table 9.3*). Even though extensive expression trials have been performed on both MBP-hYAP and His^6-hYAP, screening different bacterial strain, culture medium, incubation times, temperature and inductor concentrations, the target protein resulted invariantly localized in the insoluble fraction. Attempting to extract the protein from the inclusion bodies or to use different expression systems will be evaluated in future studies on this target. It is worth noting that no previous reports are available in literature on the production and isolation of this target.

Sample	Studied protein	Sequences
hYAP full lenght	MBP-hYAP full lenght	(-392)MKIKTGARILALSALTTMMFSASALAKIEEGKLVI WINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEK FPQVAATGDGPDIIFWAHDRFGGYAQSGLLAEITPDKAF QDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPN PPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIA ADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLI KNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNID TSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKE LAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELV KDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINA ASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGIEGRIS HMSMGGRDIVDGSLVPR↓GSEF (1)DPGQQPPPQPAPQGQGQPPSQPPQGQGPPSGPGQPAP AATQAAPQAPPAGHQIVHVRGDSETDLEALFNAVMNPK TANVPQTVPMRLRKLPDSFFKPPEPKSHSRQASTDAGTA GALTPQHVRAHSSPASLQLGAVSPGTLTPTGVVSGPAATP TAQHLRQSSFEIPDDVPLPAGWEMAKTSSGQRYFLNHID QTTTWQDPRKAMLSQMNVTAPTSPPVQQNMMNSASGP LPDGWEQAMTQDGEIYYINHKNKTTSWLDPRLDPRFAM NQRISQSAPVKQPPPLAPQSPQGGVMGGSNSNQQQQMR LQQLQMEKERLRLKQQELLRQAMRNINPSTANSPKCQEL ALRSQLPTLEQDGGTQNPVSSPGMSQELRTMTTNSSDPFL NSGTYHSRDESTDSGLSMSSYSVPRTPDDFLNSVDEMDTG DTINQSTLPSQQNRFPDYLEAIPGTNVDLGTLEGDGMNIEG EELMPSLQEALSSDILNDMESVLAATKLDKESFLTWL(504)
hYAP full lenght	His^6-hYAP full lenght	(-21)MGSSHHHHHHSQDPLVPR↓GS (1)EFDPGQQPPPQPAPQGQGQPPSQPPQGQGPPSGPGQPAPA ATQAAPQAPPAGHQIVHVRGDSETDLEALFNAVMNPKTANV QTVPMRLRKLPDSFFKPPEPKSHSRQASTDAGTAGALTPQHV RAHSSPASLQLGAVSPGTLTPTGVVSGPAATPTAQHLRQSSFEI PDDVPLPAGWEMAKTSSGQRYFLNHIDQTTTWQDPRKAMSQ MNVTAPTSPPVQQNMMNSASGPLPDGWEQAMTQDGEIYYIN HKNKTTSWLDPRLDPRFAMNQRISQSAPVKQPPPLAPQSPGGV MGGSNSNQQQQMRLQQLQMEKERLRLKQQELLRQMRNINPS TANSPKCQELALRSQLPTLEQDGGTQNPVSSPGMSQELRTMTT NSSDPFLNSGTYHSRDESTDSGLSMSSYSVPRTPDDFLNSVDE MDTGDTINQSTLPSQQNRFPDYLEAIPGTNVDLGTLEGDGMNI EGEELMPSLQEALSSDILNDMESVLAATKLDKESFLTWL(504)

Table 9.3: Summary of the investigated MBP-full length hYAP and His^6-full length hYAP. Amino acids belonging to the His6-tag and to the thrombin recognition site are coloured green and red, respectively, while the sequences of hYAP1 full length is in black.

9.3.2 Co-expression, co-purification and characterization of the XL complex

The *hYAP* gene was sub-cloned into the MCS1 of pRSFDuet-1 expression vector, already including the *hTEAD4 YBD* the MCS2 (*Figure 8.5*). The resulting expression vector allows the co-production of hYAP, bearing a thrombin cleavable N-terminal His^6 tag, and of hTEAD4 YBD (protein sequences are reported in *Table 9.4*). At variance with the isolated hYAP that was expressed in the insoluble cellular fraction in *E. coli*, the His^6-hYAP:hTEAD4 YBD XL complex was mainly produced as soluble proteins in the *E. coli* ArcticExpress (DE3) strain. The expression protocol was the same adopted to produce the His^6-hYAP TBD:hTEAD4 YBD S-L complexes. The purification of the XL complex relied on a first stage of nickel affinity chromatography (*Figure 9.26*), exploiting the N-terminal His^6-tag present on hYAP. The presence of the target proteins in the eluted fractions was verified by SDS-PAGE and WB analyses (*Figure 9.26 B* and *C*).

Sample	**Studied protein-protein complex**	**aminoacidic Sequence**
Complex XL	hYAP1 (1-504)	(-21)MGSSHHHHHHSQDPLVPR↓ GSA (1)GHQIVHVRGDSETDLEALFNAVMNPKTANVPQT VPMRLRKLPDSFFKPPEPKSHSRQASTDAGTAGALTP(70)
	hTEAD4 (217-434)	(1)MRSVASSKLWMLEFSAFLEQQQDPDTYNKHLFVH IGQSSPSYSDPYLEAVDIRQIYDKFPEKKGGLKDLFER GPSNAFFLVKFWADLNTNIEDEGSSFYGVSSQYESPEN MIITCSTKVCSFGKQVVEKVETEYARYENGHYSYRIHR SPLCEYMINFIHKLKHLPEKYMMNSVLENFTILQVVTN RDTQETLLCIAYVFEVSASEHGAQHHIYRLVKE(219)

Table 9.4: Summary of the investigated XL complex. Amino acids belonging to the His^6-tag and to the thrombin recognition site are colored green and red, respectively, while the sequences of hTEAD4 YBD and hYAP1 full length are in black.

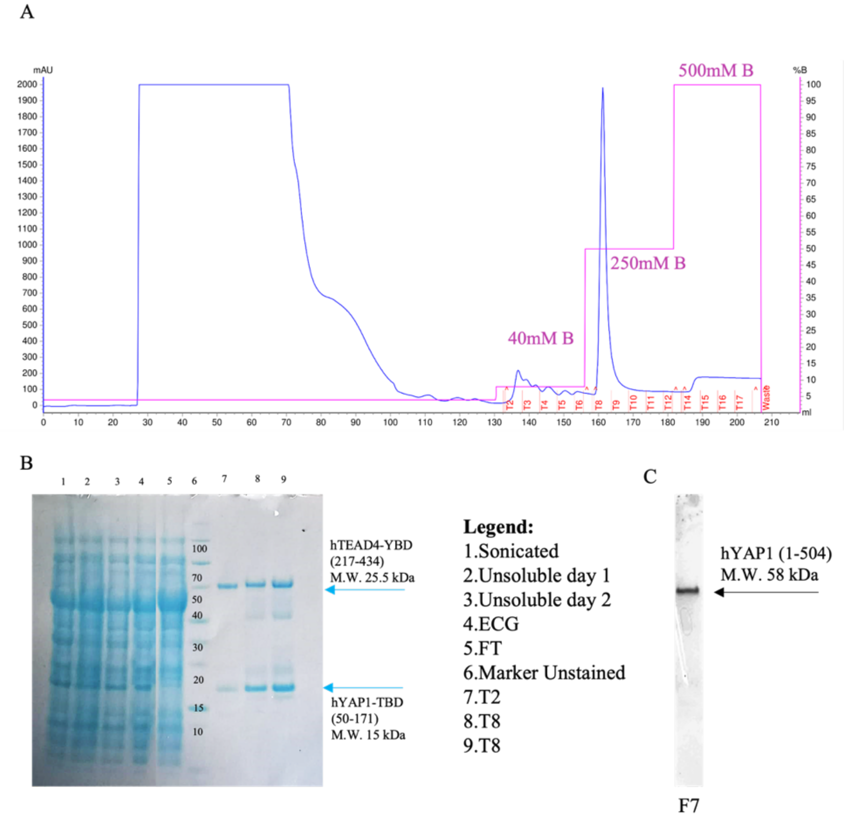


Figure 9.26: (**A**) Chromatogram of the complex XL purification through Ni^{2+} affinity chromatography (HisTrap FF 5mL column, GE-Healthcare). The UV_{280nm} monitoring and imidazole concentration are shown as blue and pink lines, respectively, whereas; the elution fractions are indicated by red marks on the x-axis. (**B**) SDS-PAGE analysis of the fractions collected from the first stage of Ni^{2+} affinity chromatography. The protein ladder is in lane 7. (**C**) Western-blot analysis on the fractions collected from Ni^{2+} affinity chromatography. The detection of recombinant His^6-tagged proteins is achieve using an HRP-conjugated antibody able to recognize the His^6 epitope.

In the second purification stage the different aggregated states present in solution were separated according to their molecular weight by means of SEC on an HiLoad 16/600 Superdex 200pg column (*Figure 9.27*). The elution profile obtained from this second purification stage evidenced by the separation of four species eluted at diverse retention volumes, according to their molecular weights. The SDS-PAGE analysis of the eluted fractions showed the presence of both His^6-hYAP and hTEAD4 YBD only in two main fractions, F5 and F3, having retention volumes compatible with the molecular weights of the XL complex and an aggregate formed by two XL complexes (2x XL complex) (*Figure 9.27*).

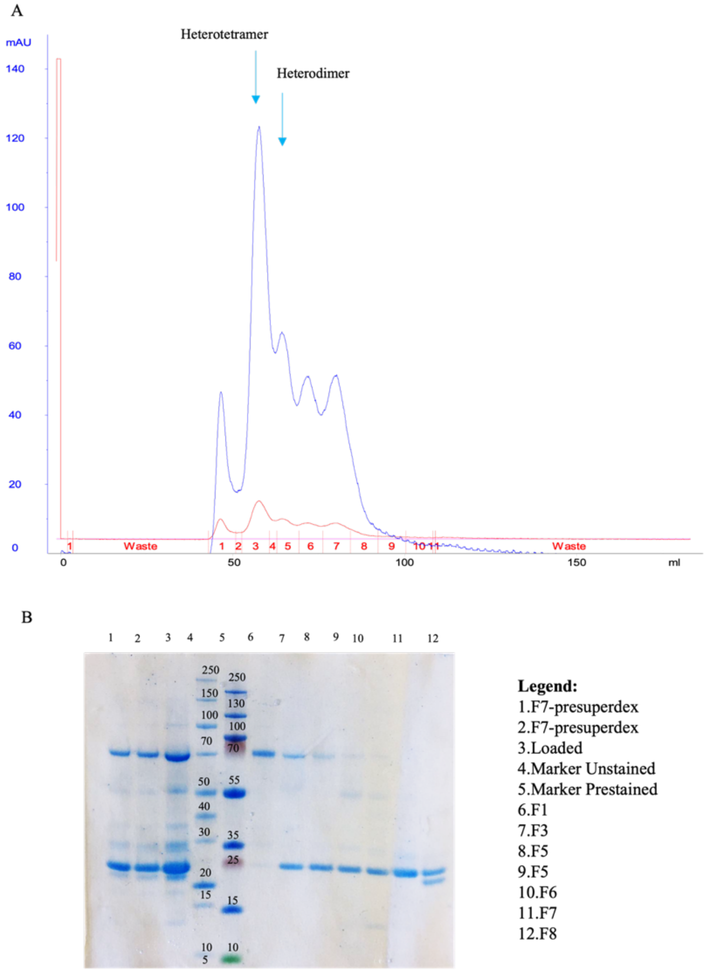
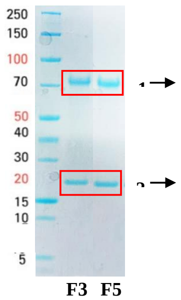


Figure 9.27: (**A**) Chromatogram of the HiLoad 16/600 Superdex 200pg column (GE-Healthcare) purification step. The UV$_{280nm}$ monitoring of protein elution is in blue; fractions are indicated by red marks on the x-axis. (**B**) SDS-PAGE analysis of the fractions collected from the SEC. The protein ladders are in lanes 5 and 6.

The presence of His^6-hYAP and hTEAD4 YBD in these fractions was also confirmed by peptide Mass Fingerprinting analysis on the bands separated by SDS-PAGE (*Figure 9.28 A*) (analysis performed on service by Dr. Laura Salvini at the TLS facility). The mass spectra of the peptide mixtures resulting from the trypsin digestion of each band, for both F3 and F5 purified fractions, resulted in the unambiguous identification of both proteins (*Figure 9.28*). For band 1 and band 2 the coverage is equal to 93.3% and 97.3%, respectively (*Figure 9.28 A*).

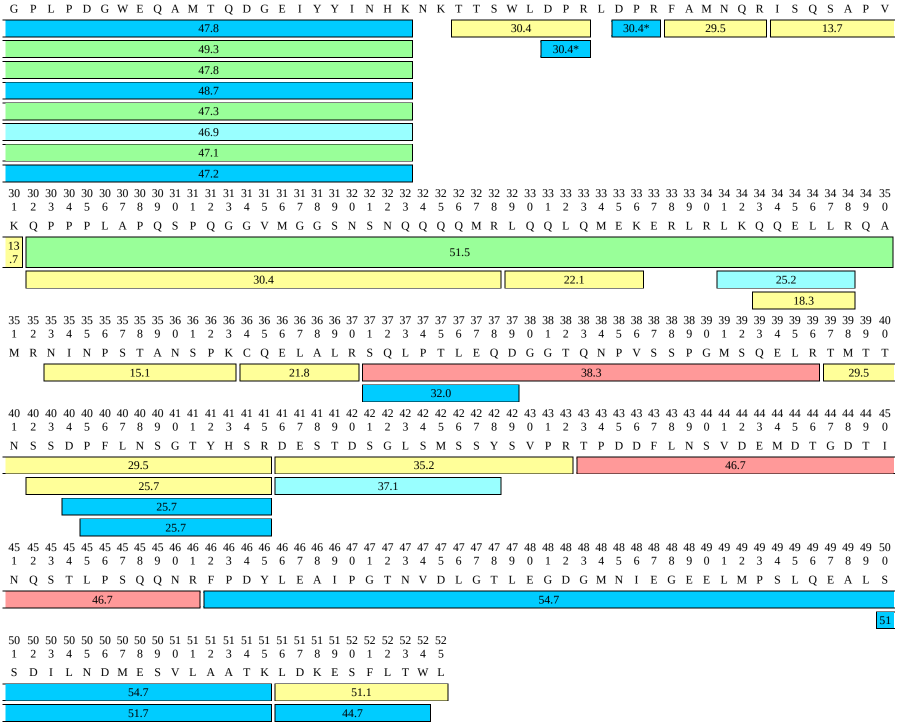
A

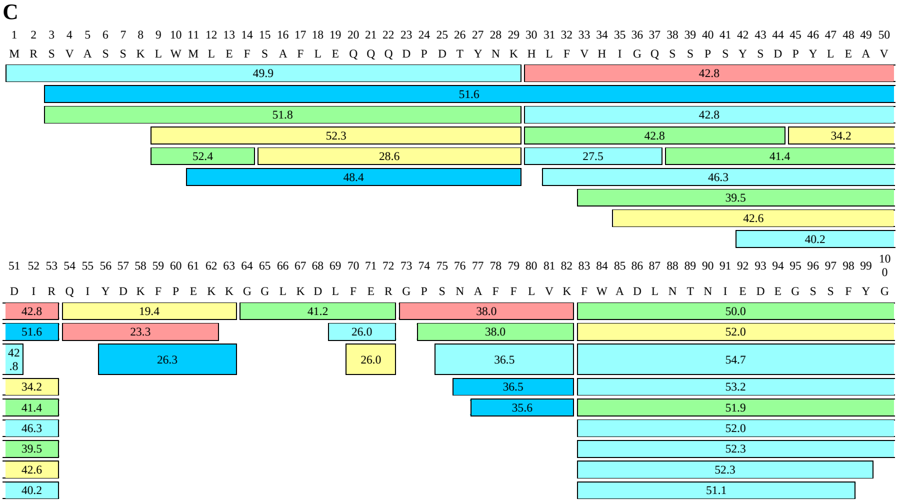


Proteins	Number of MS Peaks	Sequence Coverage
1	397	93.3%

Proteins	Number of MS Peaks	Sequence Coverage
2	785	97.3%

B


G P L P D G W E Q A M T Q D G E I Y Y I N H K N K T T S W L D P R L D P R F A M N Q R I S Q S A P V
47.8
30.4
30.4*
29.5
13.7
49.3
30.4*
47.8
48.7
47.3
46.9
47.1
47.2
30 30 30 30 30 30 30 30 30 31 31 31 31 31 31 31 31 31 31 32 32 32 32 32 32 32 32 32 32 33 33 33 33 33 33 33 33 33 33 34 34 34 34 34 34 34 34 34 34 35
1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
K Q P P P L A P Q S P Q G G V M G G S N S N Q Q Q Q M R L Q Q L Q M E K E R L R L K Q Q E L L R Q A
13 .7
51.5
30.4
22.1
25.2
18.3
35 35 35 35 35 35 35 35 35 36 36 36 36 36 36 36 36 36 36 37 37 37 37 37 37 37 37 37 37 38 38 38 38 38 38 38 38 38 38 39 39 39 39 39 39 39 39 39 39 40
1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
M R N I N P S T A N S P K C Q E L A L R S Q L P T L E Q D G G T Q N P V S S P G M S Q E L R T M T T
15.1
21.8
38.3
29.5
32.0
40 40 40 40 40 40 40 40 40 41 41 41 41 41 41 41 41 41 41 42 42 42 42 42 42 42 42 42 42 43 43 43 43 43 43 43 43 43 43 44 44 44 44 44 44 44 44 44 44 45
1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
N S S D P F L N S G T Y H S R D E S T D S G L S M S S Y S V P R T P D D F L N S V D E M D T G D T I
29.5
35.2
46.7
25.7
37.1
25.7
25.7
45 45 45 45 45 45 45 45 45 46 46 46 46 46 46 46 46 46 46 47 47 47 47 47 47 47 47 47 47 48 48 48 48 48 48 48 48 48 48 49 49 49 49 49 49 49 49 49 49 50
1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
N Q S T L P S Q Q N R F P D Y L E A I P G T N V D L G T L E G D G M N I E G E E L M P S L Q E A L S
46.7
54.7
51
50 50 50 50 50 50 50 50 50 51 51 51 51 51 51 51 51 51 51 52 52 52 52 52
1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
S D I L N D M E S V L A A T K L D K E S F L T W L
54.7
51.1
51.7
44.7


C
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
M R S V A S S S K L W M L E F S A F L E Q Q Q D P D T Y N K H L F V H I G Q S S P S Y S D P Y L E A V
49.9
42.8
51.6
51.8
42.8
52.3
42.8
34.2
52.4
28.6
27.5
41.4
48.4
46.3
39.5
42.6
40.2
51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 10 0
D I R Q I Y D K F P E K K G G L K D L F E R G P S N A F F L V K F W A D L N T N I E D E G S S F Y G
42.8
19.4
41.2
38.0
50.0
51.6
23.3
26.0
38.0
52.0
42 .8
26.3
26.0
36.5
54.7
34.2
36.5
53.2
41.4
35.6
51.9
46.3
52.0
39.5
52.3
42.6
52.3
40.2
51.1

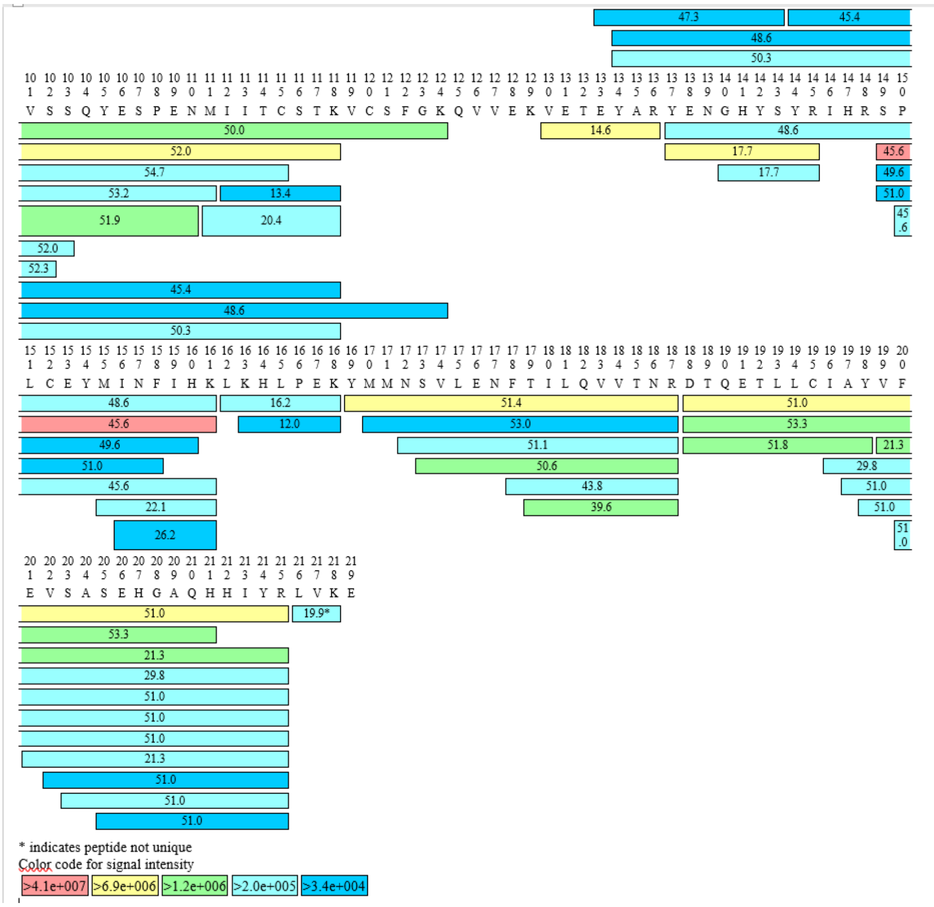


Figure 9.28: (**A**) SDS PAGE analysis for hYAP:hTEAD4 YBD XL complex F3 and F5. Schematic representation of the identified peptides by peptide mass fingerprinting after trypsin digestion of band 1 (**B**) and band 2 (**C**). The peptides mixture is correlated to the sequence of hYAP (Uniprot id P46937) and hTEAD4 (Uniprot id Q15561). The legend with the colour code for signal intensity is also shown.

The two fractions separated by SEC and attributed as the XL complex and the aggregate 2x XL complex were then subjected to native PAGE analysis (*Figure 9.29*). At variance with SDS-PAGE where proteins are denaturated, losing their secondary, tertiary, and quaternary structures, in native PAGE protein structures are retained during the analysis, including the formation of macromolecular complexes. The results of the native PAGE analysis performed on these samples confirmed the presence of the His^6-hYAP:hTEAD4 YBD XL complex in fraction F5 and of the aggregate 2x XL complex in fraction F3 (*Figure 9.29*). At variance with the S and L complex resulting monodisperse in the heterodimeric form in solution (*section 9.2.1*), different aggregated states have been observed and separated for the XL complex.

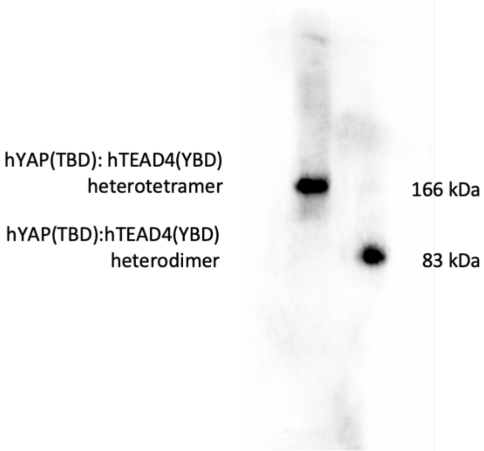
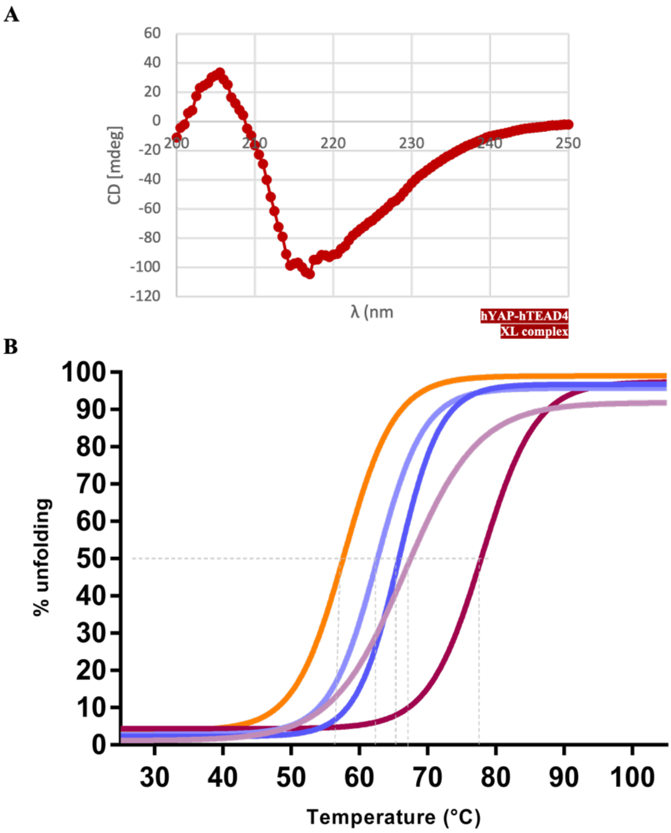


Figure 9.29: Native-PAGE, WB analysis of fractions F3 and F5 of complex XL, His6 – hYAP:hTEAD4 YBD, separated by SEC. The migration of the single band of fraction F3 corresponds to a MW of 166 kDa, indicating the presence of an heterotetrametric assembly in solution. On the other hand, the migration of the single band of fraction F5 corresponds to a MW of 83 kDa, indicating presence of His6 – hYAP:hTEAD4 YBD heterodimers in solution. In the Western-blot analysis, the detection of His^6-tagged proteins is performed using an HRP-conjugated anti His^6-tag antibody.

Afterwards, the thermal stability of the XL complex was determined by CD thermal denaturation analysis resulting in a $\mathrm{T_m}$ value of 78.01 (±0.1) °C (dark red unfolding transition curve in *Figure 9.30 A*). The unfolding transition curve obtained for the XL complex was then compared with those formerly determined for the isolated hTEAD4 YBD and for the S-L complexes, reporting an interesting trend in the $\mathrm{T_m}$ values. Indeed upon formation of the XL complex a $\mathrm{T_m}$ increment of ≈22 °C is observed with respect to the isolated hTEAD4 YBD (*Figure 9.30 B*). Notably, the $\mathrm{T_m}$ of the XL complex is also ≈12 °C higher than that of the L complex. Thus, the $\mathrm{T_m}$ increment observed by increasing the length of the hYAP TBD fragment employed in complex with hTEAD4 YBD (thus from the S to the L complex) is further extended by the interaction with full-length hYAP (complex XL). This finding has prompted the idea that other interaction areas between hYAP and hTEAD4 YBD, further the known TBD, may exist. This hypothesis is consistent also with the data achieved through SEC and native-PAGE analyses, whereby the presence of different aggregate forms was found in solution. To our knowledge, this is the first report of isolation and characterization of the hYAP:hTEAD4 YBD complex.



Protein sample	Melting Temperature (Tm)°C
hTEAD4-YBD	57.82 ± 0.1
hYAP-hTEAD4 S complex	62.55 ± 0.1
hYAP-hTEAD4 M complex	65.73 ± 0.1
hYAP-hTEAD4 L complex	66.68 ± 0.2
hYAP-hTEAD4 XL complex	78.01 ± 0.1

Figure 9.30: In the upper panel (**A**) the CD curve for XL complex. In the lower panel circular dichroism spectrum of complexes S,M,L and XL. In the lower panel (**B**), thermal unfolding transition curves of hTEAD4 YBD and of complexes S,M,L and XL followed by circular dichroism. Measurements were performed in triplicate. Melting temperature (T_m) values determined for the transitions is reported in the table.

9.3.3 Preliminary structural studies of full length hYAP in complex with hTEAD4 YBD

Preliminary attempting to unveil the structure of full legth hYAP in complex with hTEAD4 YBD were performed by combing different techniques, X-ray crystallography, Bio-SAXS and cryo-EM. As first instance, the crystallization of the XL complex heterodimer was attempted by screening $\approx$ 100 different crystallization conditions. Unfortunately, to date, no crystal growth has been observed in these trials. The limited amount of XL complex heterodimer, separated by SEC, prevented us by screening further crystallization conditions. Future, extended crystallization trials are needed to attempt the structural characterization of this complex by means of X-ray crystallography. As second instance, preliminarily Bio-SAXS experiments have been performed on both the XL complex and the aggregate 2xXL complex to achieve information on the protein folding, flexibility, and ordering of our complex. Different protein concentrations were screened both in batch or in SEC-coupled methods. The results of these analyses showed the presence of multiple aggregate forms in solution, preventing further characterizations by the *in-batch* method. Better results have been achieved by SAXS analyses combined with SEC, carried out for the samples of the XL complex and of the aggregate 2xXL complex at the concentrations of 6 mg mL^{-1} and 12 mg mL^{-1}. The first result achieved by these analyses was the observation of a concentration-dependent aggregation of the complex. Indeed, in the high-concentration (12 mg mL^{-1}) sample protein aggregates are predominant, whereas in the low-concentration (6 mg mL^{-1}) sample the presence of isolated heterodimeric XL complexes is mainly observed. Furthermore, these preliminary Bio-SAXS data suggest that both proteins involved in the XL complex have an overall fold but hYAP has some unstructured, flexible parts. Deeper Bio-SAXS investigations are needed to provide a more detailed structural characterization of the target complex. Finally, we also performed preliminary cryo-EM studies on the sample of both the heterodimer and the aggregate 2xXL complex at the FloCEN Cryo-EM facility. The sample of the aggregate 2xXL complex resulted non homogenous and it was excluded from the analysis. On the other hand, the sample of the heterodimeric XL complex resulted more homogeneous, but the observed single particle had limited dimensions to allow their clustering and structural reconstruction. More potent microscopes are needed to attempt the visualization of the heterodimeric XL complex. A major issue that can prevent the structural characterization of this complex by all these techniques is the presence of disordered, flexible regions. Indeed, it has been widely reported in literature that hYAP is an intrinsically disordered protein, which assumes a three-dimensional folding only upon binding to a partner protein [190]. Even so, no structural information is available on the isolated full length hYAP and on its complexes with partner proteins.

Chapter 10

Conclusion

hYAP and hTEAD4 proteins are downstream effectors of the Hippo pathway and their interaction is essential for the expression of genes involved in cell apoptosis, proliferation, survival and growth processes [164]. Alterations of the hYAP:hTEAD4 complex and of its transcriptional activity contribute to the insurgence and development of various hallmarks of cancer cell, making the targeting of this PPI an attractive therapeutic strategy [44, 149, 164]. One fascinating matter in cancer biology is the deeper understanding of the interaction mechanisms occurring among oncoproteins widely involved in several types of cancers, such as hYAP and hTEADs, aiming to develop novel strategies for cancer treatment. As matter of fact, the development of new drugs able to modulate the hYAP1:hTEADs interaction could represent a new powerful weapon to fight cancer, but this is still an open challenge [44, 173, 176]. In this investigation, we have established novel reliable protocols for expression and purification of the isolated hTEAD4 YBD and for the co-expression and co-purification of its complexes with different hYAP1 TBD fragments and the full length hYAP. Here, hTEAD4 YBD and its S, M, L and XL complexes were investigated by CD thermal denaturation analyses, showing meaningful increments in the thermal stability of the complexes with respect to the isolated hTEAD4 YBD. Notably, the trend of this increment in the complexes correlates with hYAP length; indeed, the XL complex, employing full length hYAP, has the highest T_m value. Even though, the structural characterization of the hYAP:hTEAD4 complex is still a challenge. We attempted the structural characterization of this complex by different techniques, X-ray crystallography, bio-SAXS and cryo-EM, providing only preliminary data that deserve deeper investigations. On the hand, we successed in the determination of the X-ray crystallographic structure of the complex S, expanding the current structural information available on this PPI. Indeed, by means of the achieved results we have demonstrated the formation of three interface regions between hYAP1 TBD and hTEAD4 YBD, with the resolution of the interface 1. Taken together, our results contribute to expand the current knowledge on the interaction between hYAP and hTEAD4, providing interesting new clues exploitable for the rational design and the development of molecules able to modulate this PPI playing a critical role in various types of cancer.

Bibliography

[1] Castrense Savojardo, Pier Luigi Martelli, and Rita Casadio. “Protein–Protein Interaction Methods and Protein Phase Separation”. In: *Annual Review of Biomedical Data Science* 3.1 (2020), pp. 89–112. DOI: `10.1146/annurev-biodatasci-011720-104428`.

[2] Teresa M. Przytycka and Dong-Yeon Cho. “Interactome”. In: John Wiley & Sons, Ltd, 2012. ISBN: 978-3-527-60090-8. DOI: `10.1002/3527600906.mcb.201100018`.

[3] William Garland, Robert Benezra, and Jaideep Chaudhary. “Chapter Fifteen - Targeting Protein–Protein Interactions to Treat Cancer—Recent Progress and Future Directions”. In: *Annual Reports in Medicinal Chemistry*. Vol. 48. Academic Press, Jan. 1, 2013, pp. 227–245. DOI: `10.1016/B978-0-12-417150-3.00015-6`.

[4] Xiaoqing Peng et al. “Protein–protein interactions: detection, reliability assessment and applications”. In: *Briefings in Bioinformatics* 18.5 (July 2016), pp. 798–819. ISSN: 1467-5463. DOI: `10.1093/bib/bbw066`.

[5] Tord Berggård, Sara Linse, and Peter James. “Methods for the detection and analysis of protein–protein interactions”. In: *PROTEOMICS* 7.16 (2007). ISSN: 1615-9861. DOI: `10.1002/pmic.200700131`.

[6] V. Srinivasa Rao et al. “Protein-protein interaction detection: methods and analysis”. In: *International Journal of Proteomics* 2014 (2014), p. 147648. ISSN: 2090-2166. DOI: `10.1155/2014/147648`.

[7] E M Phizicky and S Fields. “Protein-protein interactions: methods for detection and analysis.” In: *Microbiological Reviews* 59.1 (Mar. 1995), pp. 94–123. ISSN: 0146-0749.

[8] Mayssam H. Ali and Barbara Imperiali. “Protein oligomerization: How and why”. In: *Bioorganic & Medicinal Chemistry.* Symposium-in-Print: Chemistry and Biology of Natural Products 13.17 (Sept. 1, 2005), pp. 5013–5020. ISSN: 0968-0896. DOI: `10.1016/j.bmc.2005.05.037`.

[9] J. Monod, J. Wyman, and J. P. Changeux. “on the nature of allosteric transitions: a plausible model”. In: *Journal of Molecular Biology* 12 (May 1965), pp. 88–118. ISSN: 0022-2836. DOI: `10.1016/s0022-2836(65)80285-6`.

[10] Irene M.A. Nooren and Janet M. Thornton. “Diversity of protein–protein interactions”. In: *The EMBO Journal* 22.14 (2003), pp. 3486–3492. DOI: `https://doi.org/10.1093/emboj/cdg359`.

[11] Elisa Martino et al. “Mapping, Structure and Modulation of PPI”. In: *Frontiers in Chemistry* 9 (2021), p. 843. ISSN: 2296-2646. DOI: `10.3389/fchem.2021.718405`.

[12] Buyong Ma et al. “Protein–protein interactions: Structurally conserved residues distinguish between binding sites and exposed protein surfaces”. In: *Proceedings of the National Academy of Sciences* 100.10 (2003), pp. 5772–5777. ISSN: 0027-8424. DOI: `10.1073/pnas.1030237100`.

[13] Andrei A. Ivanov, Fadlo R. Khuri, and Haian Fu. “Targeting protein–protein interactions as an anticancer strategy”. In: *Trends in Pharmacological Sciences* 34.7 (July 1, 2013). Publisher: Elsevier, pp. 393–400. ISSN: 0165-6147. DOI: `10.1016/j.tips.2013.04.007`.

[14] K. S. Thorn and A. A. Bogan. “ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions”. In: *Bioinformatics (Oxford, England)* 17.3 (Mar. 2001), pp. 284–285. ISSN: 1367-4803. DOI: `10.1093/bioinformatics/17.3.284`.

[15] Engin Cukuroglu et al. “Hot spots in protein–protein interfaces: Towards drug discovery”. In: *Progress in Biophysics and Molecular Biology* 116.2 (2014), pp. 165–173. ISSN: 0079-6107. DOI: `https://doi.org/10.1016/j.pbiomolbio.2014.06.003`. URL: `https://www.sciencedirect.com/science/article/pii/S0079610714000455`.

[16] K. L. Morrison and G. A. Weiss. “Combinatorial alanine-scanning”. In: *Current Opinion in Chemical Biology* 5.3 (June 2001), pp. 302–307. ISSN: 1367-5931. DOI: `10.1016/s1367-5931(00)00206-4`.

[17] Harry C. Jubb et al. “Mutations at protein-protein interfaces: Small changes over big surfaces have large impacts on human health”. In: *Progress in Biophysics and Molecular Biology* 128 (2017). Exploring mechanisms in biology: simulations and experiments come together, pp. 3–13. ISSN: 0079-6107. DOI: `https://doi.org/10.1016/j.pbiomolbio.2016.10.002`.

[18] Hannes Ponstingl et al. “Morphological aspects of oligomeric protein structures”. In: *Progress in Biophysics and Molecular Biology* 89.1 (Sept. 2005), pp. 9–35. ISSN: 0079-6107. DOI: `10.1016/j.pbiomolbio.2004.07.010`.

[19] Chung-Jung Tsai et al. “Studies of protein-protein interfaces: A statistical analysis of the hydrophobic effect”. In: *Protein Science* 6.1 (1997), pp. 53–64. ISSN: 1469-896X. DOI: `10.1002/pro.5560060106`.

[20] J. Brange and L. Langkjoer. “Insulin structure and stability”. In: *Pharmaceutical Biotechnology* 5 (1993), pp. 315–350. ISSN: 1078-0467. DOI: `10.1007/978-1-4899-1236-7_11`.

[21] Adam M. Koppel and Jonathan A. Raper. “Collapsin-1 Covalently Dimerizes, and Dimerization Is Necessary for Collapsing Activity *”. In: *Journal of Biological Chemistry* 273.25 (June 19, 1998), pp. 15708–15713. ISSN: 0021-9258, 1083-351X. DOI: `10.1074/jbc.273.25.15708`. (Visited on 01/11/2022).

[22] Mads Østergaard, Narendra Kumar Mishra, and Knud J. Jensen. "The ABC of Insulin: The Organic Chemistry of a Small Protein". In: *Chemistry – A European Journal* 26.38 (2020), pp. 8341–8357. ISSN: 1521-3765. DOI: `10.1002/chem.202000337`.

[23] Mainak Guharoy and Pinak Chakrabarti. "Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein–protein interactions". In: *Bioinformatics* 23.15 (Aug. 1, 2007), pp. 1909–1918. ISSN: 1367-4803. DOI: `10.1093/bioinformatics/btm274`.

[24] Katja Luck et al. "A reference map of the human binary protein interactome". In: *Nature* 580.7803 (Apr. 2020), pp. 402–408. ISSN: 1476-4687. DOI: `10.1038/s41586-020-2188-x`.

[25] Daniel P. Ryan and Jacqueline M. Matthews. "Protein-protein interactions in human disease". In: *Current Opinion in Structural Biology* 15.4 (Aug. 2005), pp. 441–446. ISSN: 0959-440X. DOI: `10.1016/j.sbi.2005.06.001`.

[26] Haiying Lu et al. "Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials". In: *Signal Transduction and Targeted Therapy* 5, 213, 1 (Sept. 2020). ISSN: 2059-3635. DOI: `10.1038/s41392-020-00315-3`.

[27] Andrea G.S. Buggins and Chris J. Pepper. "The role of Bcl-2 family proteins in chronic lymphocytic leukaemia". In: *Leukemia Research* 34.7 (2010), pp. 837–842. ISSN: 0145-2126. DOI: `https://doi.org/10.1016/j.leukres.2010.03.011`.

[28] Tracy L. Nero et al. "Oncogenic protein interfaces: small molecules, big challenges". In: *Nature Reviews. Cancer* 14.4 (Apr. 2014), pp. 248–262. ISSN: 1474-1768. DOI: `10.1038/nrc3690`.

[29] Jesus Perez de Vega M., Martin-Martinez Mercedes, and Gonzalez-Muniz Rosario. "Modulation of Protein-Protein Interactions by Stabilizing/Mimicking Protein Secondary Structure Elements". In: *Current Topics in Medicinal Chemistry* 7.1 (Dec. 31, 2006), pp. 33–62.

[30] Ji-Wei Chang et al. “Prediction of Protein–Protein Interactions by Evidence Combining Methods”. In: *International Journal of Molecular Sciences* 17.11 (Nov. 2016), p. 1946. DOI: `10.3390/ijms17111946`.

[31] Robert B Russell et al. “A structural perspective on protein–protein interactions”. In: *Current Opinion in Structural Biology* 14.3 (June 1, 2004), pp. 313–324. ISSN: 0959-440X. DOI: `10.1016/j.sbi.2004.04.006`.

[32] Eric Ennifar. “X-ray crystallography as a tool for mechanism-of-action studies and drug discovery”. In: *Current Pharmaceutical Biotechnology* 14.5 (2013), pp. 537–550. ISSN: 1873-4316. DOI: `10.2174/13892010140513111104824`.

[33] Adam Ben-Shem et al. “Crystal Structure of the Eukaryotic Ribosome”. In: *Science* (Nov. 26, 2010). DOI: `10.1126/science.1194294`.

[34] V. Ramakrishnan. “Ribosome structure and the mechanism of translation”. In: *Cell* 108.4 (Feb. 22, 2002), pp. 557–572. ISSN: 0092-8674. DOI: `10.1016/s0092-8674(02)00619-0`.

[35] M. M. Yusupov et al. “Crystal structure of the ribosome at 5.5 A resolution”. In: *Science (New York, N.Y.)* 292.5518 (May 4, 2001), pp. 883–896. ISSN: 0036-8075. DOI: `10.1126/science.1060089`.

[36] Gulnara Yusupova and Marat Yusupov. “Crystal structure of eukaryotic ribosome and its complexes with inhibitors”. In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 372.1716 (Mar. 19, 2017), p. 20160184. ISSN: 1471-2970. DOI: `10.1098/rstb.2016.0184`.

[37] Wayne Harshbarger et al. “Crystal Structure of the Human 20S Proteasome in Complex with Carfilzomib”. In: *Structure* 23.2 (Feb. 2015), pp. 418–424. ISSN: 09692126. DOI: `10.1016/j.str.2014.11.017`.

[38] Masaki Unno et al. “The structure of the mammalian 20S proteasome at 2.75 A resolution”. In: *Structure (London, England: 1993)* 10.5 (May 2002), pp. 609–618. ISSN: 0969-2126. DOI: `10.1016/s0969-2126(02)00748-7`.

[39] Andreas Schweitzer et al. “Structure of the human 26S proteasome at a resolution of 3.9 Å”. In: *Proceedings of the National Academy of Sciences* 113.28 (2016), pp. 7816–7821. ISSN: 0027-8424. DOI: `10.1073/pnas.1608050113`.

[40] Ewan Phillip Ramsay et al. “Structure of human RNA polymerase III”. In: *Nature Communications* 11.1 (2020), p. 6409. ISSN: 2041-1723. DOI: `10.1038/s41467-020-20262-5`.

[41] Dong Wang et al. “X-ray structure and mechanism of RNA polymerase II stalled at an antineoplastic monofunctional platinum-DNA adduct”. In: *Proceedings of the National Academy of Sciences* 107.21 (2010), pp. 9584–9589. ISSN: 0027-8424. DOI: `10.1073/pnas.1002565107`.

[42] Andrey G. Baranovskiy et al. “X-ray structure of the complex of regulatory subunits of human DNA polymerase delta”. In: *Cell Cycle* 7.19 (2008), pp. 3026–3036. DOI: `10.4161/cc.7.19.6720`.

[43] Laura Taddia et al. “Inside the biochemical pathways of thymidylate synthase perturbed by anticancer drugs: Novel strategies to overcome cancer chemoresistance”. In: *Drug Resistance Updates* 23 (Nov. 1, 2015), pp. 20–54. ISSN: 1368-7646. DOI: `10.1016/j.drup.2015.10.003`.

[44] Matteo Santucci et al. “The Hippo Pathway and YAP/TAZ-TEAD Protein-Protein Interaction as Targets for Regenerative Medicine and Cancer Treatment”. In: *Journal of Medicinal Chemistry* 58.12 (June 25, 2015), pp. 4857–4873. ISSN: 1520-4804. DOI: `10.1021/jm501615v`.

[45] Divita Garg et al. “Novel approaches for targeting thymidylate synthase to overcome the resistance and toxicity of anticancer drugs”. In: *Journal of Medicinal Chemistry* 53.18 (Sept. 23, 2010), pp. 6539–6549. ISSN: 1520-4804. DOI: `10.1021/jm901869w`.

[46] C. W. Carreras and D. V. Santi. “The catalytic mechanism and structure of thymidylate synthase”. In: *Annual Review of Biochemistry* 64 (1995), pp. 721–762. ISSN: 0066-4154. DOI: `10.1146/annurev.bi.64.070195.003445`.

[47] Stefania Ferrari et al. “Human Thymidylate Synthase Inhibitors Halting Ovarian Cancer Growth”. In: *Vitamins and Hormones* 107 (2018), pp. 473–513. ISSN: 0083-6729. DOI: `10.1016/bs.vh.2017.12.002`.

[48] Michal G. Rose, Michael P. Farrell, and John C. Schmitz. “Thymidylate synthase: a critical target for cancer chemotherapy”. In: *Clinical Colorectal Cancer* 1.4 (Feb. 2002), pp. 220–229. ISSN: 1533-0028. DOI: `10.3816/CCC.2002.n.003`.

[49] Franklin G. Berger and Sondra H. Berger. “Thymidylate synthase as a chemotherapeutic drug target: Where are we after fifty years?” In: *Cancer Biology & Therapy* 5.9 (Sept. 2006), pp. 1238–1241. ISSN: 1538-4047, 1555-8576. DOI: `10.4161/cbt.5.9.3414`.

[50] Ningwen Tai et al. “Translational autoregulation of thymidylate synthase and dihydrofolate reductase”. In: *FBL* 9.4 (2004), pp. 2521–2526. DOI: `10.2741/1413`.

[51] Svetlana A. Kholodar and Amnon Kohen. “Noncovalent Intermediate of Thymidylate Synthase: Fact or Fiction?” In: *Journal of the American Chemical Society* 138.26 (July 6, 2016). Publisher: American Chemical Society, pp. 8056–8059. ISSN: 0002-7863. DOI: `10.1021/jacs.6b03826`.

[52] Natalia Kanaan et al. “A Quantum Mechanics/Molecular Mechanics Study of the Catalytic Mechanism of the Thymidylate Synthase”. In: *Biochemistry* 46.12 (Mar. 1, 2007). Publisher: American Chemical Society, pp. 3704–3713. ISSN: 0006-2960. DOI: `10.1021/bi061953y`.

[53] Robert M. Stroud and Janet S. Finer-Moore. “Conformational Dynamics along an Enzymatic Reaction Pathway: Thymidylate Synthase, “the Movie””. In: *Biochemistry* 42.2 (Jan. 1, 2003). Publisher: American Chemical Society, pp. 239–247. ISSN: 0006-2960. DOI: `10.1021/bi020598i`.

[54] Q. Cui et al. “Thymidylate kinase: an old topic brings new perspectives”. In: *Current Medicinal Chemistry* 20.10 (2013), pp. 1286–1305. ISSN: 1875-533X. DOI: `10.2174/0929867311320100006`.

[55] Peter J. Houghton. “Thymineless Death”. In: *Antifolate Drugs in Cancer Therapy.* Ed. by Ann L. Jackman. Cancer Drug Discovery and Development. Totowa, NJ: Humana Press, 1999, pp. 423–435. ISBN: 978-1-59259-725-3. DOI: `10.1007/978-1-59259-725-3_21`.

[56] J. A. Houghton, F. G. Harwood, and P. J. Houghton. “Cell cycle control processes determine cytostasis or cytotoxicity in thymineless death of colon cancer cells”. In: *Cancer Research* 54.18 (Sept. 15, 1994), pp. 4967–4973. ISSN: 0008-5472.

[57] B. Kamen. “Folate and antifolate pharmacology”. In: *Seminars in Oncology* 24.5 (Oct. 1997), S18–30–S18–39. ISSN: 0093-7754.

[58] Daniela Cardinale et al. “Protein–protein interface-binding peptides inhibit the cancer therapy target human thymidylate synthase”. In: *Proceedings of the National Academy of Sciences* 108.34 (Aug. 23, 2011), E542–E549. ISSN: 0027-8424, 1091-6490. DOI: `10.1073/pnas.1104829108`.

[59] P. Deschamps et al. “Crystal structure of the active form of native human thymidylate synthase in the absence of bound substrates”. In: *Acta Crystallographica. Section F, Structural Biology Communications* 73 (Pt 6 May 25, 2017), pp. 336–341. ISSN: 2053-230X. DOI: `10.1107/S2053230X17007233`.

[60] Janet S. Finer-Moore, Daniel V. Santi, and Robert M. Stroud. “Lessons and conclusions from dissecting the mechanism of a bisubstrate enzyme: thymidylate synthase mutagenesis, function, and structure”. In: *Biochemistry* 42.2 (Jan. 21, 2003), pp. 248–256. ISSN: 0006-2960. DOI: `10.1021/bi020599a`.

[61] Rami Almog et al. “Crystal structure of a deletion mutant of human thymidylate synthase D (7–29) and its ternary complex with Tomudex and dUMP”. In: *Protein Science : A Publication of the Protein Society* 10.5 (May 2001), pp. 988–996. ISSN: 0961-8368.

[62] Cecilia Pozzi et al. “Structural Bases for the Synergistic Inhibition of Human Thymidylate Synthase and Ovarian Cancer Cell Growth by Drug Combinations”. In: *Cancers* 13.9 (Jan. 2021), p. 2061. DOI: `10.3390/cancers13092061`.

[63] Filippo Genovese et al. “Dimer–monomer equilibrium of human thymidylate synthase monitored by fluorescence resonance energy transfer”. In: *Protein Science : A Publication of the Protein Society* 19.5 (May 2010), pp. 1023–1030. ISSN: 0961-8368. DOI: `10.1002/pro.379`.

[64] Jason Phan et al. “Structure of Human Thymidylate Synthase Suggests Advantages of Chemotherapy with Noncompetitive Inhibitors *”. In: *Journal of Biological Chemistry* 276.17 (Apr. 27, 2001). Publisher: Elsevier, pp. 14170–14177. ISSN: 0021-9258, 1083-351X. DOI: `10.1074/jbc.M009493200`.

[65] Leslie L Lovelace et al. “Variants of human thymidylate synthase with loop 181–197 stabilized in the inactive conformation”. In: *Protein Science : A Publication of the Protein Society* 18.8 (Aug. 2009), pp. 1628–1636. ISSN: 0961-8368. DOI: `10.1002/pro.171`. (Visited on 01/21/2022).

[66] Lydia M. Gibson, Leslie L. Lovelace, and Lukasz Lebioda. “The R163K Mutant of Human Thymidylate Synthase Is Stabilized in an Active Conformation: Structural Asymmetry and Reactivity of Cysteine 195”. In: *Biochemistry* 47.16 (Apr. 22, 2008), pp. 4636–4643. ISSN: 0006-2960. DOI: `10.1021/bi7019386`.

[67] Dan Chen et al. “Structural analyses of human thymidylate synthase reveal a site that may control conformational switching between active and inactive states”. In: *The Journal of Biological Chemistry* 292.32 (Aug. 11, 2017), pp. 13449–13458. ISSN: 1083-351X. DOI: `10.1074/jbc.M117.787267`.

[68] Thomas J. Stout and Robert M. Stroud. “The complex of the anti-cancer therapeutic, BW1843U89, with thymidylate synthase at 2.0 å resolution: implications for a new mode of inhibition”. In: *Structure* 4.1 (Jan. 1, 1996).

Publisher: Elsevier, pp. 67–77. ISSN: 0969-2126. DOI: `10.1016/S0969-2126(96)00010-X`.

[69] Xiao Huang et al. "Replacement of Val3 in Human Thymidylate Synthase Affects Its Kinetic Properties and Intracellular Stability," in: *Biochemistry* 49.11 (Mar. 23, 2010). Publisher: American Chemical Society, pp. 2475–2482. ISSN: 0006-2960. DOI: `10.1021/bi901457e`.

[70] Outi M. H. Salo-Ahen et al. "Hotspots in an Obligate Homodimeric Anticancer Target. Structural and Functional Effects of Interfacial Mutations in Human Thymidylate Synthase". In: *Journal of medicinal chemistry* 58.8 (Apr. 23, 2015), pp. 3572–3581. ISSN: 0022-2623. DOI: `10.1021/acs.jmedchem.5b00137`.

[71] Anna Tochowicz et al. "Alanine Mutants of the Interface Residues of Human Thymidylate Synthase Decode Key Features of the Binding Mode of Allosteric Anticancer Peptides". In: *Journal of Medicinal Chemistry* 58.2 (2015), pp. 1012–1018. DOI: `10.1021/jm5011176`.

[72] Mireia Rosell and Juan Fernández-Recio. "Hot-spot analysis for drug discovery targeting protein-protein interactions". In: *Expert Opinion on Drug Discovery* 13.4 (Apr. 3, 2018), pp. 327–338. ISSN: 1746-0441, 1746-045X. DOI: `10.1080/17460441.2018.1430763`.

[73] Ruth L. Saxl et al. "Significance of mutations on the structural perturbation of thymidylate synthase: Implications for their involvement in subunit exchange". In: *Protein Science : A Publication of the Protein Society* 16.7 (July 2007), pp. 1439–1448. ISSN: 0961-8368. DOI: `10.1110/ps.062509807`.

[74] Jun Liu et al. "Thymidylate synthase as a translational regulator of cellular gene expression". In: *Biochimica Et Biophysica Acta* 1587.2 (July 18, 2002), pp. 174–182. ISSN: 0006-3002. DOI: `10.1016/s0925-4439(02)00080-7`.

[75] Edward Chu and Carmen J. Allegra. "The role of thymidylate synthase as an RNA binding protein". In: *BioEssays* 18.3 (1996). _eprint:

https://onlinelibrary.wiley.com/doi/pdf/10.1002/bies.950180306, pp. 191–198. ISSN: 1521-1878. DOI: `10.1002/bies.950180306`.

[76] Donna M. Voeller et al. “The identification of thymidylate synthase peptide domains located in the interface region that bind thymidylate synthase mRNA”. In: *Biochemical and Biophysical Research Communications* 297.1 (Sept. 13, 2002), pp. 24–31. ISSN: 0006-291X. DOI: `10.1016/s0006-291x(02)02080-6`.

[77] Nicholas D. Brunn et al. “Analysis of mRNA recognition by human thymidylate synthase”. In: *Bioscience Reports* 34.6 (Dec. 23, 2014), e00168. ISSN: 0144-8463. DOI: `10.1042/BSR20140137`.

[78] E. Chu et al. “Autoregulation of human thymidylate synthase messenger RNA translation by thymidylate synthase.” In: *Proceedings of the National Academy of Sciences* 88.20 (Oct. 15, 1991), pp. 8977–8981. ISSN: 0027-8424, 1091-6490. DOI: `10.1073/pnas.88.20.8977`.

[79] Abul Hasnat et al. “Fluorescence Spectroscopic and 19F NMR Studies of Human Thymidylate Synthase with its Cognate RNA”. In: *Journal of Biomolecular Structure and Dynamics* 25.3 (Dec. 1, 2007), pp. 253–269. ISSN: 0739-1102. DOI: `10.1080/07391102.2007.10507174`.

[80] X. Lin et al. “In vitro selection of an RNA sequence that interacts with high affinity with thymidylate synthase”. In: *Nucleic Acids Research* 28.21 (Nov. 1, 2000), pp. 4266–4274. ISSN: 1362-4962. DOI: `10.1093/nar/28.21.4266`.

[81] Xiukun Lin et al. “Role of cysteine amino acid residues on the RNA binding activity of human thymidylate synthase”. In: *Nucleic Acids Research* 31.16 (Aug. 15, 2003), pp. 4882–4887. ISSN: 0305-1048. DOI: `10.1093/nar/gkg678`.

[82] K. Keyomarsi et al. “The thymidylate synthase inhibitor, ICI D1694, overcomes translational detainment of the enzyme”. In: *The Journal of Biological Chemistry* 268.20 (July 15, 1993), pp. 15142–15149. ISSN: 0021-9258.

[83] E. Chu et al. “Identification of an RNA binding site for human thymidylate synthase”. In: *Proceedings of the National Academy of Sciences of the United States of America* 90.2 (1993), pp. 517–521. ISSN: 0027-8424. DOI: `10.1073/pnas.90.2.517`.

[84] Sondra H. Berger, Franklin G. Berger, and Lukasz Lebioda. “Effects of ligand binding and conformational switching on intracellular stability of human thymidylate synthase”. In: *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1696.1 (Jan. 14, 2004), pp. 15–22. ISSN: 1570-9639. DOI: `10.1016/j.bbapap.2003.09.005`.

[85] Peter M. Wilson et al. “Standing the test of time: targeting thymidylate biosynthesis in cancer therapy”. In: *Nature Reviews Clinical Oncology* 11.5 (2014), pp. 282–298. ISSN: 1759-4782. DOI: `10.1038/nrclinonc.2014.51`. URL: `https://doi.org/10.1038/nrclinonc.2014.51`.

[86] D. Papamichael. “The Use of Thymidylate Synthase Inhibitors in the Treatment of Advanced Colorectal Cancer: Current Status”. In: *STEM CELLS* 18.3 (2000). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1634/stemcells.18-3-166, pp. 166–175. ISSN: 1549-4918. DOI: `10.1634/stemcells.18-3-166`.

[87] Daniel B. Longley, D. Paul Harkin, and Patrick G. Johnston. “5-Fluorouracil: mechanisms of action and clinical strategies”. In: *Nature Reviews Cancer* 3.5 (2003), pp. 330–338. ISSN: 1474-1768. DOI: `10.1038/nrc1074`. URL: `https://doi.org/10.1038/nrc1074`.

[88] Giuseppe Aprile et al. “Pharmacology and therapeutic efficacy of capecitabine: focus on breast and colorectal cancer”. In: *Anti-Cancer Drugs* 20.4 (2009). ISSN: 0959-4973.

[89] Kazumasa Ikeda et al. “Bioactivation of Tegafur to 5-Fluorouracil Is Catalyzed by Cytochrome P-450 2A6 in Human Liver Microsomes in Vitro1”. In: *Clinical Cancer Research* 6.11 (Nov. 2000), pp. 4409–4415. ISSN: 1078-0432.

[90] Yehuda G. Assaraf. “Molecular basis of antifolate resistance”. In: *Cancer Metastasis Reviews* 26.1 (Mar. 2007), pp. 153–181. ISSN: 0167-7659. DOI: `10.1007/s10555-007-9049-z`.

[91] G. R. Westerhof et al. “Carrier- and receptor-mediated transport of folate antagonists targeting folate-dependent enzymes: correlates of molecular-structure and biological activity”. In: *Molecular Pharmacology* 48.3 (Sept. 1995), pp. 459–471. ISSN: 0026-895X.

[92] Michele Visentin et al. “The membrane transport and polyglutamation of pralatrexate: a new-generation dihydrofolate reductase inhibitor”. In: *Cancer Chemotherapy and Pharmacology* 72.3 (Sept. 2013), pp. 597–606. ISSN: 1432-0843. DOI: `10.1007/s00280-013-2231-9`.

[93] Esti Liani et al. “Loss of folylpoly-g-glutamate synthetase activity is a dominant mechanism of resistance to polyglutamylation-dependent novel antifolates in multiple human leukemia sublines”. In: *International Journal of Cancer* 103.5 (2003), pp. 587–599. DOI: `https://doi.org/10.1002/ijc.10829`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/ijc.10829`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1002/ijc.10829`.

[94] Vittorio D. Ferrari et al. “Epirubicin, cisplatin, and raltitrexed in patients with advanced gastric and hepatobiliary carcinoma: a phase II study”. In: *American Journal of Clinical Oncology* 27.5 (Oct. 2004), pp. 445–448. ISSN: 1537-453X. DOI: `10.1097/01.coc.0000128722.41587.19`.

[95] A. Planting et al. “Phase I study of concomitant chemoradiation with raltitrexed in locally advanced head and neck cancer”. In: *European Journal of Cancer (Oxford, England: 1990)* 41.1 (Jan. 2005), pp. 93–97. ISSN: 0959-8049. DOI: `10.1016/j.ejca.2004.09.022`.

[96] Jan P. van Meerbeeck et al. “Randomized phase III study of cisplatin with or without raltitrexed in patients with malignant pleural mesothelioma: an intergroup study of the European Organisation for Research and Treatment of Cancer Lung Cancer Group and the National

Cancer Institute of Canada". In: *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 23.28 (Oct. 1, 2005), pp. 6881–6889. ISSN: 0732-183X. DOI: `10.1200/JCO.20005.14.589`.

[97] G. K. Schwartz, A. Harstrick, and M. González Barón. "Raltitrexed (Tomudex) in combination with 5-fluorouracil for the treatment of patients with advanced colorectal cancer: preliminary results from phase I clinical trials". In: *European Journal of Cancer (Oxford, England: 1990)* 35 Suppl 1 (Mar. 1999), S9–13. ISSN: 0959-8049. DOI: `10.1016/s0959-8049(99)00042-8`.

[98] Aleem Gangjee, Hiteshkumar D. Jain, and Sonali Kurup. "Recent advances in classical and non-classical antifolates as antitumor and antiopportunistic infection agents: part I". In: *Anti-Cancer Agents in Medicinal Chemistry* 7.5 (Sept. 2007), pp. 524–542. ISSN: 1871-5206. DOI: `10.2174/187152007781668724`.

[99] ANN L. Jackman, MARTIN Forster, and MATTHEW Ng. "CHAPTER 9 - Targeting thymidylate synthase by antifolate drugs for the treatment of cancer". In: *Cancer Drug Design and Discovery.* Ed. by Stephen Neidle. New York: Academic Press, Jan. 1, 2008, pp. 198–226. ISBN: 978-0-12-369448-5. DOI: `10.1016/B978-012369448-5.50012-4`.

[100] S. Webber et al. "AG337, a novel lipophilic thymidylate synthase inhibitor: in vitro and in vivo preclinical studies". In: *Cancer Chemotherapy and Pharmacology* 37.6 (1996), pp. 509–517. ISSN: 0344-5704. DOI: `10.1007/s002800050422`.

[101] Y. Takemura, H. Kobayashi, and H. Miyachi. "Cellular and molecular mechanisms of resistance to antifolate drugs: new analogues and approaches to overcome the resistance". In: *International Journal of Hematology* 66.4 (Dec. 1997), pp. 459–477. ISSN: 0925-5710. DOI: `10.1016/s0925-5710(97)00058-3`.

[102] Yuzo Sato et al. “Upregulation of Thymidylate Synthase Induces Pemetrexed Resistance in Malignant Pleural Mesothelioma”. In: *Frontiers in Pharmacology* 12 (2021). ISSN: 1663-9812.

[103] G. Jansen and R. Pieters. “The role of impaired transport in (pre)clinical resistance to methotrexate: insights on new antifolates”. In: *Drug Resistance Updates: Reviews and Commentaries in Antimicrobial and Anticancer Chemotherapy* 1.3 (1998), pp. 211–218. ISSN: 1368-7646. DOI: `10.1016/s1368-7646(98)80042-3`.

[104] Gerrit Jansen et al. “A Structurally Altered Human Reduced Folate Carrier with Increased Folic Acid Transport Mediates a Novel Mechanism of Antifolate Resistance*”. In: *Journal of Biological Chemistry* 273.46 (Nov. 13, 1998), pp. 30189–30198. ISSN: 0021-9258. DOI: `10.1074/jbc.273.46.30189`.

[105] P H Sayre et al. “Multi-targeted antifolates aimed at avoiding drug resistance form covalent closed inhibitory complexes with human and Escherichia coli thymidylate synthases”. In: *Journal of molecular biology* 313.4 (Nov. 1, 2001), pp. 813–829. ISSN: 1089-8638. DOI: `10.1006/jmbi.2001.5074`.

[106] Alessio Ligabue et al. “Transcriptional activation and cell cycle block are the keys for 5-fluorouracil induced up-regulation of human thymidylate synthase expression”. In: *PloS One* 7.10 (2012), e47318. ISSN: 1932-6203. DOI: `10.1371/journal.pone.0047318`.

[107] Lei Zheng, Ulrich Baumann, and Jean-Louis Reymond. “An efficient one-step site-directed and site-saturation mutagenesis protocol”. In: *Nucleic Acids Research* 32.14 (Aug. 10, 2004), e115. ISSN: 1362-4962. DOI: `10.1093/nar/gnh110`.

[108] Cecilia Pozzi et al. “Evidence of Destabilization of the Human Thymidylate Synthase (hTS) Dimeric Structure Induced by the Interface Mutation Q62R”. In: *Biomolecules* 9.4 (Apr. 3, 2019), E134. ISSN: 2218-273X. DOI: `10.3390/biom9040134`.

[109] Manuela Benvenuti and Stefano Mangani. "Crystallization of soluble proteins in vapor diffusion for x-ray crystallography". In: *Nature Protocols* 2.7 (July 2007). Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7 Primary_atype: Protocols Publisher: Nature Publishing Group, pp. 1633–1651. ISSN: 1750-2799. DOI: `10.1038/nprot.2007.198`.

[110] W. Kabsch. "XDS". In: *Acta Crystallographica Section D: Biological Crystallography* 66.2 (Feb. 1, 2010). Publisher: International Union of Crystallography, pp. 125–132. ISSN: 0907-4449. DOI: `10.1107/S0907444909047337`.

[111] Philip Evans. "Scaling and assessment of data quality". In: *Acta Crystallographica. Section D, Biological Crystallography* 62 (Pt 1 Jan. 2006), pp. 72–82. ISSN: 0907-4449. DOI: `10.1107/S0907444905036693`.

[112] Philip R. Evans. "An introduction to data reduction: space-group determination, scaling and intensity statistics". In: *Acta Crystallographica. Section D, Biological Crystallography* 67 (Pt 4 Apr. 2011), pp. 282–292. ISSN: 1399-0047. DOI: `10.1107/S090744491003982X`.

[113] Martyn D. Winn et al. "Overview of the CCP4 suite and current developments". In: *Acta Crystallographica. Section D, Biological Crystallography* 67 (Pt 4 Apr. 2011), pp. 235–242. ISSN: 1399-0047. DOI: `10.1107/S0907444910045749`.

[114] Alexei Vagin and Alexei Teplyakov. "Molecular replacement with MOLREP". In: *Acta Crystallographica. Section D, Biological Crystallography* 66 (Pt 1 Jan. 2010), pp. 22–25. ISSN: 1399-0047. DOI: `10.1107/S0907444909042589`.

[115] Garib N. Murshudov et al. "REFMAC5 for the refinement of macromolecular crystal structures". In: *Acta Crystallographica. Section D, Biological Crystallography* 67 (Pt 4 Apr. 2011), pp. 355–367. ISSN: 1399-0047. DOI: `10.1107/S0907444911001314`.

[116] P. Emsley et al. “Features and development of Coot”. In: *Acta Crystallographica. Section D, Biological Crystallography* 66 (Pt 4 Apr. 2010), pp. 486–501. ISSN: 1399-0047. DOI: `10.1107/S0907444910007493`.

[117] R. A. Laskowski et al. “PROCHECK: a program to check the stereochemical quality of protein structures”. In: *Journal of Applied Crystallography* 26.2 (1993). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1107/S0021889892009944, pp. 283–291. ISSN: 1600-5767. DOI: `10.1107/S0021889892009944`.

[118] Giacomo Janson et al. “PyMod 2.0: improvements in protein sequence-structure analysis and homology modeling within PyMOL”. In: *Bioinformatics* 33.3 (Oct. 2016), pp. 444–446. ISSN: 1367-4803. DOI: `10.1093/bioinformatics/btw638`.

[119] S. McNicholas et al. “Presenting your structures: the CCP4mg molecular-graphics software”. In: *Acta Crystallographica. Section D, Biological Crystallography* 67 (Pt 4 Apr. 2011), pp. 386–394. ISSN: 1399-0047. DOI: `10.1107/S0907444911007281`.

[120] Norma J. Greenfield. “Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions”. In: *Nature Protocols* 1.6 (Dec. 2006). Bandiera_abtest: a Cg_type: Nature Research Journals Number: 6 Primary_atype: Protocols Publisher: Nature Publishing Group, pp. 2527–2535. ISSN: 1750-2799. DOI: `10.1038/nprot.2006.204`.

[121] Patricia S. Kumagai, Ana P. U. Araujo, and Jose L. S. Lopes. “Going deep into protein secondary structure with synchrotron radiation circular dichroism spectroscopy”. In: *Biophysical Reviews* 9.5 (Aug. 19, 2017), pp. 517–527. ISSN: 1867-2450. DOI: `10.1007/s12551-017-0314-2`.

[122] Larry W. Hardy et al. “Atomic Structure of Thymidylate Synthase: Target for Rational Drug Design”. In: *Science* (Jan. 23, 1987). Publisher: American Association for the Advancement of Science. DOI: `10.1126/science.3099389`.

[123] Celia A. Schiffer et al. “Crystal structure of human thymidylate synthase: a structural mechanism for guiding substrates into the active site”. In: *Biochemistry* 34.50 (Dec. 19, 1995). Publisher: American Chemical Society, pp. 16279–16287. ISSN: 0006-2960. DOI: `10.1021/bi00050a007`.

[124] F. William Studier. “Protein production by auto-induction in high density shaking cultures”. In: *Protein Expression and Purification* 41.1 (May 2005), pp. 207–234. ISSN: 1046-5928. DOI: `10.1016/j.pep.2005.01.016`.

[125] Glauco Ponterini et al. “Intracellular quantitative detection of human thymidylate synthase engagement with an unconventional inhibitor using tetracysteine-diarsenical-probe technology”. In: *Scientific Reports* 6.1 (June 2, 2016). Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group, p. 27198. ISSN: 2045-2322. DOI: `10.1038/srep27198`.

[126] J. Phan et al. “Human thymidylate synthase is in the closed conformation when complexed with dUMP and raltitrexed, an antifolate drug”. In: *Biochemistry* 40.7 (Feb. 20, 2001), pp. 1897–1902. ISSN: 0006-2960. DOI: `10.1021/bi002413i`.

[127] Evgeny Krissinel and Kim Henrick. “Inference of macromolecular assemblies from crystalline state”. In: *Journal of Molecular Biology* 372.3 (Sept. 21, 2007), pp. 774–797. ISSN: 0022-2836. DOI: `10.1016/j.jmb.2007.05.022`.

[128] P. Stover and V. Schirch. “The metabolic role of leucovorin”. In: *Trends in Biochemical Sciences* 18.3 (Mar. 1993), pp. 102–106. ISSN: 0968-0004. DOI: `10.1016/0968-0004(93)90162-g`.

[129] Cecilia Pozzi et al. “Structural and Functional Characterization of the Human Thymidylate Synthase (hTS) Interface Variant R175C, New Perspectives for the Development of hTS Inhibitors”. In: *Molecules*

(Basel, Switzerland) 24.7 (Apr. 7, 2019), E1362. ISSN: 1420-3049. DOI: `10.3390/molecules24071362`.

[130] Cecilia Pozzi et al. "Structural Comparison of Enterococcus faecalis and Human Thymidylate Synthase Complexes with the Substrate dUMP and Its Analogue FdUMP Provides Hints about Enzyme Conformational Variabilities". In: *Molecules (Basel, Switzerland)* 24.7 (Mar. 31, 2019), E1257. ISSN: 1420-3049. DOI: `10.3390/molecules24071257`.

[131] Jacalyn M. Green and Rowena G. Matthews. "Folate Biosynthesis, Reduction, and Polyglutamylation and the Interconversion of Folate Derivatives". In: *EcoSal Plus* (Mar. 21, 2007). Publisher: ASM PressWashington, DC. DOI: `10.1128/ecosalplus.3.6.3.6`.

[132] Helena Almqvist et al. "CETSA screening identifies known and novel thymidylate synthase inhibitors and slow intracellular activation of 5-fluorouracil". In: *Nature Communications* 7.1 (Mar. 24, 2016). Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Molecular biophysics;Screening;Target validation Subject_term_id: molecular-biophysics;screening;target-validation, p. 11040. ISSN: 2041-1723. DOI: `10.1038/ncomms11040`. URL: `https://www.nature.com/articles/ncomms11040` (visited on 01/21/2022).

[133] Hannu Myllykallio et al. "An alternative flavin-dependent mechanism for thymidylate synthesis". In: *Science (New York, N.Y.)* 297.5578 (July 5, 2002), pp. 105–107. ISSN: 1095-9203. DOI: `10.1126/science.1072113`.

[134] Tatiana V. Mishanina et al. "An unprecedented mechanism of nucleotide methylation in organisms containing thyX". In: *Science (New York, N.Y.)* 351.6272 (Jan. 29, 2016), pp. 507–510. ISSN: 1095-9203. DOI: `10.1126/science.aad0300`.

[135] Fakhira Hassan Nazki, Aga Syed Sameer, and Bashir Ahmad Ganaie. “Folate: metabolism, genes, polymorphisms and the associated diseases”. In: *Gene* 533.1 (Jan. 1, 2014), pp. 11–20. ISSN: 1879-0038. DOI: `10.1016/j.gene.2013.09.063`.

[136] Paul J. Sapienza, Bradley T. Falk, and Andrew L. Lee. “Bacterial Thymidylate Synthase Binds Two Molecules of Substrate and Cofactor without Cooperativity”. In: *Journal of the American Chemical Society* 137.45 (Nov. 18, 2015), pp. 14260–14263. ISSN: 1520-5126. DOI: `10.1021/jacs.5b10128`.

[137] Anatoly Chernyshev, Todd Fleischmann, and Amnon Kohen. “Thymidyl biosynthesis enzymes as antibiotic targets”. In: *Applied Microbiology and Biotechnology* 74.2 (Feb. 2007), pp. 282–289. ISSN: 0175-7598. DOI: `10.1007/s00253-006-0763-1`.

[138] Michael Choi, Kalani Karunaratne, and Amnon Kohen. “Flavin-Dependent Thymidylate Synthase as a New Antibiotic Target”. In: *Molecules (Basel, Switzerland)* 21.5 (May 20, 2016), E654. ISSN: 1420-3049. DOI: `10.3390/molecules21050654`.

[139] Randy Johnson and Georg Halder. “The two faces of Hippo: targeting the Hippo pathway for regenerative medicine and cancer treatment”. In: *Nature Reviews Drug Discovery* 13.1 (2014), pp. 63–79. ISSN: 1474-1784. DOI: `10.1038/nrd4161`.

[140] Bin Zhao, Karen Tumaneng, and Kun-Liang Guan. “The Hippo pathway in organ size control, tissue regeneration and stem cell self-renewal”. In: *Nature Cell Biology* 13.8 (2011), pp. 877–883. ISSN: 1476-4679. DOI: `10.1038/ncb2303`.

[141] Shenghong Ma et al. “The Hippo Pathway: Biology and Pathophysiology”. In: *Annual Review of Biochemistry* 88.1 (2019). _eprint: https://doi.org/10.1146/annurev-biochem-013118-111829, pp. 577–604. DOI: `10.1146/annurev-biochem-013118-111829`.

[142] Jyoti R. Misra and Kenneth D. Irvine. “The Hippo Signaling Network and Its Biological Functions”. In: *Annual Review of Genetics* 52.1 (2018). _eprint: https://doi.org/10.1146/annurev-genet-120417-031621, pp. 65–87. DOI: `10.1146/annurev-genet-120417-031621`. URL: `https://doi.org/10.1146/annurev-genet-120417-031621` (visited on 12/27/2021).

[143] Kieran Harvey and Nicolas Tapon. “The Salvador–Warts–Hippo pathway — an emerging tumour-suppressor network”. In: *Nature Reviews Cancer* 7.3 (Mar. 2007), pp. 182–191. ISSN: 1474-1768. DOI: `10.1038/nrc2070`.

[144] Chu Zhu, Li Li, and Bin Zhao. “The regulation and function of YAP transcription co-activator”. In: *Acta Biochimica Et Biophysica Sinica* 47.1 (Jan. 2015), pp. 16–28. ISSN: 1745-7270. DOI: `10.1093/abbs/gmu110`.

[145] Boon Chin Heng et al. “An overview of signaling pathways regulating YAP/TAZ activity”. In: *Cellular and Molecular Life Sciences* 78.2 (Jan. 1, 2021), pp. 497–512. ISSN: 1420-9071. DOI: `10.1007/s00018-020-03579-8`.

[146] Floriane Gibault et al. “Targeting Transcriptional Enhanced Associate Domains (TEADs)”. In: *Journal of Medicinal Chemistry* 61.12 (2018). PMID: 29251924, pp. 5057–5072. DOI: `10.1021/acs.jmedchem.7b00879`.

[147] Kimberly C. Lin, Hyun Woo Park, and Kun-Liang Guan. “Regulation of the Hippo Pathway Transcription Factor TEAD”. In: *Trends in Biochemical Sciences* 42.11 (Nov. 1, 2017). Publisher: Elsevier, pp. 862–872. DOI: `10.1016/j.tibs.2017.09.003`. URL: `https://www.cell.com/trends/biochemical-sciences/abstract/S0968-0004(17)30170-6`.

[148] Mu Chen et al. “Structural and Functional Overview of TEAD4 in Cancer Biology”. In: *OncoTargets and therapy* 13 (Oct. 6, 2020), pp. 9865–9874. ISSN: 1178-6930. DOI: `10.2147/OTT.S266649`.

[149] Philamer C. Calses et al. “Hippo Pathway in Cancer: Aberrant Regulation and Therapeutic Opportunities”. In: *Trends in Cancer* 5.5 (May 2019), pp. 297–307. ISSN: 24058033. DOI: `10.1016/j.trecan.2019.04.001`.

[150] Ajaybabu V. Pobbati and Wanjin Hong. “Emerging roles of TEAD transcription factors and its coactivators in cancers”. In: *Cancer Biology & Therapy* 14.5 (May 1, 2013), pp. 390–398. ISSN: 1538-4047. DOI: `10.4161/cbt.23788`.

[151] Makiko Fujii et al. “TGF-β synergizes with defects in the Hippo pathway to stimulate human malignant mesothelioma growth”. In: *The Journal of Experimental Medicine* 209.3 (Mar. 12, 2012), pp. 479–494. ISSN: 0022-1007. DOI: `10.1084/jem.20111653`.

[152] L. H. Chan et al. “Hedgehog signaling induces osteosarcoma development through Yap1 and H19 overexpression”. In: *Oncogene* 33.40 (Oct. 2, 2014), pp. 4857–4866. ISSN: 1476-5594. DOI: `10.1038/onc.2013.433`.

[153] Chun-Li Da et al. “Significance and relationship between Yes-associated protein and survivin expression in gastric carcinoma and precancerous lesions”. In: *World Journal of Gastroenterology : WJG* 15.32 (Aug. 28, 2009), pp. 4055–4061. ISSN: 1007-9327. DOI: `10.3748/wjg.15.4055`.

[154] A Vassilev et al. “TEAD/TEF transcription factors utilize the activation domain of YAP65, a Src/Yes-associated protein localized in the cytoplasm”. en. In: *Genes Dev.* 15.10 (May 2001), pp. 1229–1241.

[155] Jeffrey K. Holden and Christian N. Cunningham. “Targeting the Hippo Pathway and Cancer through the TEAD Family of Transcription Factors”. In: *Cancers* 10.3 (Mar. 20, 2018), p. 81. ISSN: 2072-6694. DOI: `10.3390/cancers10030081`.

[156] Yannick Mesrouze et al. “Effect of the acylation of TEAD4 on its interaction with co-activators YAP and TAZ”. In: *Protein Science : A Publication of the Protein Society* 26.12 (Dec. 2017), pp. 2399–2409. ISSN: 0961-8368. DOI: `10.1002/pro.3312`.

[157] PuiYee Chan et al. “Autopalmitoylation of TEAD proteins regulates transcriptional output of the Hippo pathway”. In: *Nature Chemical Biology* 12.4 (Apr. 2016), pp. 282–289. ISSN: 1552-4469. DOI: `10.1038/nchembio.2036`.

[158] Cameron L. Noland et al. “Palmitoylation of TEAD Transcription Factors Is Required for Their Stability and Function in Hippo Pathway Signaling”. In: *Structure* 24.1 (Jan. 2016), pp. 179–186. ISSN: 09692126. DOI: `10.1016/j.str.2015.11.005`.

[159] M. Sudol. “Yes-associated protein (YAP65) is a proline-rich phosphoprotein that binds to the SH3 domain of the Yes proto-oncogene product”. In: *Oncogene* 9.8 (Aug. 1994), pp. 2145–2152. ISSN: 0950-9232.

[160] Tsutomu Oka et al. “Functional complexes between YAP2 and ZO-2 are PDZ domain-dependent, and regulate YAP2 nuclear localization and signalling”. en. In: *Biochem. J.* 432.3 (Dec. 2010), pp. 461–472.

[161] Ze Li et al. “Structural insights into the YAP and TEAD complex”. In: *Genes & Development* 24.3 (Feb. 1, 2010), pp. 235–240. ISSN: 1549-5477. DOI: `10.1101/gad.1865810`.

[162] Marius Sudol and Kieran F. Harvey. “Modularity in the Hippo signaling pathway”. In: *Trends in Biochemical Sciences* 35.11 (Nov. 1, 2010). Publisher: Elsevier, pp. 627–633. ISSN: 0968-0004. DOI: `10.1016/j.tibs.2010.05.010`.

[163] Yannick Mesrouze et al. “Adaptation of the bound intrinsically disordered protein YAP to mutations at the YAP:TEAD interface”. In: *Protein Science: A Publication of the Protein Society* 27.10 (Oct. 2018), pp. 1810–1820. ISSN: 1469-896X. DOI: `10.1002/pro.3493`.

[164] Kieran F Harvey, Xiaomeng Zhang, and David M Thomas. “The Hippo pathway and human cancer”. In: *Nat. Rev. Cancer* 13.4 (Apr. 2013), pp. 246–257.

[165] Francesca Zanconato, Michelangelo Cordenonsi, and Stefano Piccolo. “YAP/TAZ at the Roots of Cancer”. In: *Cancer Cell* 29.6 (June 13, 2016), pp. 783–803. ISSN: 1878-3686. DOI: `10.1016/j.ccell.2016.05.005`.

[166] Francesca Cottini et al. “Rescue of Hippo coactivator YAP1 triggers DNA damage-induced apoptosis in hematological cancers”. In: *Nature Medicine* 20.6 (June 2014), pp. 599–606. ISSN: 1546-170X. DOI: `10.1038/nm.3562`.

[167] Lin Zhang et al. “The hippo pathway effector YAP regulates motility, invasion, and castration-resistant growth of prostate cancer cells”. In: *Molecular and Cellular Biology* 35.8 (Apr. 2015), pp. 1350–1362. ISSN: 1098-5549. DOI: `10.1128/MCB.00102-15`.

[168] Mohammad Reza Zinatizadeh et al. “The Hippo Tumor Suppressor Pathway (YAP/TAZ/TEAD/MST/LATS) and EGFR-RAS-RAF-MEK in cancer metastasis”. In: *Genes & Diseases* 8.1 (Jan. 1, 2021), pp. 48–60. ISSN: 2352-3042. DOI: `10.1016/j.gendis.2019.11.003`. URL: `https://www.sciencedirect.com/science/article/pii/S2352304219301163` (visited on 12/28/2021).

[169] Chunbo He et al. “The Hippo/YAP pathway interacts with EGFR signaling and HPV oncoproteins to regulate cervical cancer progression”. In: *EMBO Molecular Medicine* 7.11 (2015), pp. 1426–1449. DOI: `https://doi.org/10.15252/emmm.201404976`.

[170] Yan Hu, Hanshuo Mu, and Zhiping Deng. “The Transcription Factor TEAD4 Enhances Lung Adenocarcinoma Malignancy through Enhancing PKM2 Mediated Glycolysis”. In: *bioRxiv* (2020). DOI: `10.1101/2020.11.25.399295`.

[171] WOOJIN JEONG et al. “Activation of YAP1 Is Associated with Poor Prognosis and Response to Taxanes in Ovarian Cancer”. In: *Anticancer Research* 34.2 (2014), pp. 811–817. ISSN: 0250-7005.

[172] Yanyan Han. “Analysis of the role of the Hippo pathway in cancer”. In: *Journal of Translational Medicine* 17.1 (Apr. 8, 2019), p. 116. ISSN: 1479-5876. DOI: `10.1186/s12967-019-1869-4`.

[173] Gian Marco Elisi et al. “Repurposing of Drugs Targeting YAP-TEAD Functions”. In: *Cancers* 10.9 (Sept. 14, 2018), E329. ISSN: 2072-6694. DOI: `10.3390/cancers10090329`.

[174] Juntao Feng et al. “Verteporfin, a suppressor of YAP-TEAD complex, presents promising antitumor properties on ovarian cancer”. In: *OncoTargets and Therapy* 9 (2016), pp. 5371–5381. ISSN: 1178-6930. DOI: `10.2147/OTT.S109979`.

[175] Katarzyna Brodowska et al. “The clinically used photosensitizer Verteporfin (VP) inhibits YAP-TEAD and human retinoblastoma cell growth in vitro without light activation”. In: *Experimental eye research* 124 (July 2014), pp. 67–73. DOI: `10.1016/j.exer.2014.04.011`. URL: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4135181/`.

[176] Ajaybabu V. Pobbati and Wanjin Hong. “A combat with the YAP/TAZ-TEAD oncoproteins for cancer therapy”. In: *Theranostics* 10.8 (Feb. 18, 2020), pp. 3622–3635. ISSN: 1838-7640. DOI: `10.7150/thno.40889`.

[177] James J. Crawford, Sarah M. Bronner, and Jason R. Zbieg. “Hippo pathway inhibition by blocking the YAP/TAZ-TEAD interface: a patent review”. In: *Expert Opinion on Therapeutic Patents* 28.12 (Dec. 2018), pp. 867–873. ISSN: 1744-7674. DOI: `10.1080/13543776.2018.1549226`.

[178] Ajaybabu V. Pobbati et al. “Targeting the Central Pocket in Human Transcription Factor TEAD as a Potential Cancer Therapeutic Strategy”. In: *Structure (London, England: 1993)* 23.11 (Nov. 3, 2015), pp. 2076–2086. ISSN: 1878-4186. DOI: `10.1016/j.str.2015.09.009`.

[179] Zhisen Zhang et al. “Structure-Based Design and Synthesis of Potent Cyclic Peptides Inhibiting the YAP–TEAD Protein–Protein Interaction”. In: *ACS Medicinal Chemistry Letters* 5.9 (2014). PMID: 25221655, pp. 993–998. DOI: `10.1021/ml500160m`.

[180] Khuchtumur Bum-Erdene et al. “Small-Molecule Covalent Modification of Conserved Cysteine Leads to Allosteric Inhibition of the TEAD-Yap Protein-Protein Interaction”. In: *Cell Chemical Biology* 26.3 (Mar. 2019), 378–389.e13. ISSN: 24519456. DOI: `10.1016/j.chembiol.2018.11.010`.

[181] Terese Bergfors. “Seeds to crystals”. In: *Journal of Structural Biology* 142.1 (2003). Macromolecular crystallization in the structural genomics

era, pp. 66–76. ISSN: 1047-8477. DOI: `https://doi.org/10.1016/S1047-8477(03)00039-X`.

[182] Katherine A. Kantardjieff and Bernhard Rupp. “Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals”. In: *Protein Science : A Publication of the Protein Society* 12.9 (Sept. 2003), pp. 1865–1871. ISSN: 0961-8368.

[183] José Malanho Silva et al. “Non-crystallographic symmetry in proteins: Jahn–Teller-like and Butterfly-like effects?” In: *JBIC Journal of Biological Inorganic Chemistry* 24.1 (2019), pp. 91–101. ISSN: 1432-1327. DOI: `10.1007/s00775-018-1630-0`.

[184] R. J. Read and A. J. Schierbeek. “A phased translation function”. In: *Journal of Applied Crystallography* 21.5 (Oct. 1, 1988). Number: 5 Publisher: International Union of Crystallography, pp. 490–495. ISSN: 0021-8898. DOI: `10.1107/S002188988800562X`. URL: `//scripts.iucr.org/cgi-bin/paper?ms0244` (visited on 12/18/2021).

[185] Liz Potterton et al. “Developments in the CCP4 molecular-graphics project”. In: *Acta Crystallographica. Section D, Biological Crystallography* 60 (Pt 12 Pt 1 Dec. 2004), pp. 2288–2294. ISSN: 0907-4449. DOI: `10.1107/S0907444904023716`.

[186] Douglas R. Smyth et al. “Crystal structures of fusion proteins with large-affinity tags”. In: *Protein Science* 12.7 (2003). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1110/ps.0243403, pp. 1313–1322. ISSN: 1469-896X. DOI: `10.1110/ps.0243403`.

[187] Yannick Mesrouze et al. “Dissection of the interaction between the intrinsically disordered YAP protein and the transcription factor TEAD”. In: *eLife* 6 (2017). Ed. by Jane Clarke, e25068. ISSN: 2050-084X. DOI: `10.7554/eLife.25068`.

[188] R. T. Nolte et al. “Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma”. In: *Nature* 395.6698 (Sept. 10, 1998), pp. 137–143. ISSN: 0028-0836. DOI: `10.1038/25931`.

[189] Beatrice D. Darimont et al. “Structure and specificity of nuclear receptor–coactivator interactions”. In: *Genes & Development* 12.21 (Nov. 1, 1998), pp. 3343–3356. ISSN: 0890-9369.

[190] Wei Tian et al. “Structural and functional analysis of the YAP-binding domain of human TEAD2”. In: *Proceedings of the National Academy of Sciences of the United States of America* 107.16 (Apr. 20, 2010), pp. 7293–7298. ISSN: 1091-6490. DOI: `10.1073/pnas.1000293107`.

[191] Charles O. Rock. “CHAPTER 3 - Fatty acid and phospholipid metabolism in prokaryotes”. In: *Biochemistry of Lipids, Lipoproteins and Membranes (Fifth Edition).* Ed. by Dennis E. Vance and Jean E. Vance. San Diego: Elsevier, Jan. 1, 2008, pp. 59–96. ISBN: 978-0-444-53219-0. DOI: `10.1016/B978-044453219-0.50005-2`.

[192] M. Joseph and R. Nagaraj. “Is the role of fatty acids only to provide membrane-anchor in fatty acylated proteins?” In: *Indian Journal of Biochemistry & Biophysics* 34.1 (Apr. 1997), pp. 1–5. ISSN: 0301-1208.

Part IV

Appendices

Appendix I

List of scientific publications

- Pozzi C., **Lopresti L.**, Tassone G., Mangani S.
  Targeting Methyltransferases in Human Pathogenic Bacteria: Insights into Thymidylate Synthase (TS) and Flavin- Dependent TS (FDTS). Molecules. 2019; 24(8):1638. DOI: 10.3390/molecules24081638

- Pozzi, C., **Lopresti, L.**, Santucci, M., Costi, M.P., Mangani, S.
  Evidence of Destabilization of the Human Thymidylate Synthase (hTS) Dimeric Structure Induced by the Interface Mutation Q62R. Biomolecules 2019, 9, 134; DOI:10.3390/biom9040134.

Appendix II

Participation to conferences, congresses and schools

Oral Communication

- **Lopresti, L.**, Tassone, G., Santucci, M., Tagliazucchi, L., Costi, M.P., Pozzi, C., Mangani, S. ***Evidence of destabilization of the human Thymidylate Synthase (hTS) dimeric structure induced by the interface mutations***. Merck Young Chemists Symposium (MYCS), Rimini (Italy), 21th -24th November 2021. Proceedings of the Merck Young Chemists Symposium XX edition, page XX

- **Lopresti, L.**, Pozzi, C., Tagliazucchi, L., D'Arca, D., Marverti, G., Costi, M.P., Mangani, S. ***Structural characterization of YAP:TEAD4 protein-protein interaction, an emerging target for cancer treatment.*** XLIX Meeting of the Italian Crystallographic Association (AIC). September 6th-9th 2021, Online event

- **Lopresti L.**, Pozzi, C., Tagliazucchi, L., D'Arca, D., Marverti, G., Costi, M.P., Mangani, S. ***Structural insight into YAP-TEAD4 protein-protein interaction a target for cancer treatment.*** II Meeting of the Italian Crystallographic Association Biological Macromolecules Group. June 7th 9th, 2021, Online event.

- **Lopresti L.**, Pozzi, C., D'Arca, D., Marverti, G., Costi, M.P., Mangani, S. ***Structural insight into YAP-TEAD4 protein-protein interactions as target for cancer treatment.*** I GCI@HOME Event - Giovani Cristallografi Italiani, September 29th 2020, Online event.

Abstract & Poster Presentation

- **Lopresti L.**, Pozzi, C., Tagliazucchi, L., D'Arca, D., Marverti, G., Costi, M.P., Mangani, S. ***Structural insight into YAP-TEAD4 protein-protein interactions a target for cancer treatment.*** Poster at the XXVII National Congress of the Division of Chemistry of Biological Systems (Italian Chemical Society), SCI2021, September 14th -23rd, 2021, Online event.
- Tagliazucchi, L., Malpezzi, G., Ponterini, G., Mangani, S., **Lopresti, L.**, Pozzi, C., Venturelli, A., Costi, M.P. ***Targeting of the cysteines surface domain of the Transcription Enhancer Associate Domain (TEAD) for anticancer drug Discovery.*** Poster at the XIII Young Medicinal Chemist's Symposium - Nuove Prospettive in Chimica Farmaceutica, Young MedChem,SCI, April 20th-25th, 2021 online event.
- **Lopresti, L.**, Pozzi, C., Santucci, M., Costi, M.P., Mangani, S. ***Structural and functional characterization of new human thymidylate synthase interface variants.*** Abstract at the I Meeting of the Italian Crystallographic Association Biological Macromolecules Group. February 20th - 21st , 2021, Fiesole, Italy.
- **Lopresti, L.**, Pozzi, C., Santucci, M., Costi, M.P., Mangani, S. ***Structural and functional characterization of new human thymidylate synthase interface variants.*** Poster at the XXVI National Congress of the Division of Chemistry of Biological Systems (Italian Chemical Society), (DCSB2019), September 11th-13th, 2019, Siena, Italy.
- **Lopresti, L.**, Pozzi, C., Santucci, M., Costi, M.P., Mangani, S. ***Structural and functional characterization of new human thymidylate synthase interface variant.*** Poster at the V Conference of the Italian (AIC) and Spanish Crystallographic (GE3C) Associations (MISCA V), September 4th-7th, 2019, Naples, Italy.

School Participation

- 2nd edition of the **Instruct virtual school on Single Particle Analysis by CryoEM**, June 28th – July 2nd 2021, Online School
- **GECRY School, Session I: "From Gene to Crystal"**, January 21st - 24th, 2020, Basovizza (TS), Italy.
- **GECRY School, Session II: "From Crystal to Structure"**, September 22nd-25th, 2020, Online school.
- **PSB Symposium "Macromolecules in action"**, July 4th-5th, 2019, EPN Campus, Grenoble, France.
- **Scuola Nazionale di Chimica Bioinorganica per Dottorandi** 2019, February 12nd-15th, 2019, Rome, Italy.

Honours & Awards

- **Student participation bursary**, XXVII National Congress of the Division of Chemistry of Biological Systems (Italian Chemical Society), SCI2021, September 14th -23rd, 2021, Online event.

- **Best Poster Award** for the poster ***Structural and functional characterization of new human thymidylate synthase interface variants.*** PhD Day in Chemical and Pharmaceutical Sciences, 22nd November 2019, University of Siena, Siena, Italy.

- **Best Poster Award** for the poster ***Structural and functional characterization of new human thymidylate synthase interface variants.*** V Conference of the Italian (AIC) and Spanish Crystallographic (GE3C) Associations (MISCA V), September 4th-7th, 2019, Naples, Italy.

- **Student participation bursary**, XXVI National Congress of the Division of Chemistry of Biological Systems (Italian Chemical Society), (DCSB2019), September 11st -13rd, 2019, Siena, Italy.

- **BCC Pachino graduation award** banned by Banca di Credito Cooperativo di Pachino, May 2019, Pachino, Italy.